Functional studies of *Ci-Gsx* gene in the developing central nervous system of *Ciona intestinalis*

Thesis

Functional studies of the *Ci-gsx* gene in the developing central nervous system of *Ciona intestinalis*

A thesis submitted to the Open University of London for the degree of

DOCTOR OF PHILOSOPHY

by

Rosaria Esposito

Sponsoring Establishment:

STAZIONE ZOOLOGICA ANTON DOHRN

NAPOLI, ITALY

DATE OF SUBMISSION: 30 SEPTEMBER 2013
DATE OF AWARD: 15 JANUARY 2014

November 2013

This thesis work has been carried out in the laboratory of Dr Antonietta Spagnuolo, at the Stazione Zoologica Anton Dohrn in Napoli, Italy

Director of studies: Dr. *Antonietta Spagnuolo* (Stazione Zoologica, Napoli, Italy)

External Supervisor: Dr. *Ivan Conte* (Telethon Institute of Genetics and Medicine, Naples, Italy)

In tempi duri dobbiamo avere sogni duri, sogni reali,
quelli che, se ci daremo da fare, si avvereranno.
Clarissa Pinkola Estès

Se sono in grado di camminare da sola,
posso andare dove voglio"
Paulo Coelho

Per te Nonna.
I miei occhi non possono vederti,
ma il mio cuore sa che tu sei sempre con me.

ABSTRACT

Gsx genes encode for members of ParaHox family of transcription factor and are highly conserved throughout the animal kingdom. They are expressed mainly in the developing CNS in members of all major groups of metazoans. Gsx function, in the few organisms in which it has been investigated, is related to the specification of neural progenitors and in the regulation of the dorso-ventral patterning of the vertebrate neural tube.

In the ascidian *Ciona intestinalis*, the model organism used in this work, a single Gsx ortholog has been identified, named *Ci-gsx*, whose zygotic expression is localized in the developing CNS, including the precursors of the ocellus photoreceptor cells.

Ciona ocellus, a pigmented multicellular photoreceptive organ, despite its simplicity, possesses functional and molecular similarities to vertebrate eye, thus representing a valuable tool to study the basic mechanisms of eye development.

The aim of my PhD work has been the investigation of Ci-gsx function in the early specification of photoreceptor cells in *Ciona intestinalis*. To this end, I used targeted perturbation of Ci-gsx activity. This study indicated that Ci-gsx is actually involved in the molecular pathway controlling photoreceptor cells differentiation, upstream of *Ci-Rx*, a key gene for ocellus development. In addition, a detailed analysis of the 5' *Ci-gsx* regulatory region allowed me to identify the minimal promoter region responsible for *Ci-gsx* expression. Furthermore, I demonstrated that a member of Sox family is likely involved in *Ci-gsx* transcriptional activation. I also identified Ci-msxb as a possible repressor of *Ci-gsx* expression in pigment cell lineage.

My data insert a new piece in the genetic cascade controlling early specification of *Ciona* sensory organs and represent a first evidence of a potential implication of a Gsx ParaHox gene in this process, thus opening new perspectives on the role, still controversial, of Gsx gene in the ancestor of bilaterians.

INDEX

ABSTRACT I

INDEX II

PREFACE V

LIST OF FIGURES VI

LIST OF TABLES VIII

Chapter 1. Introduction ..... - 1 -

THE MODEL ORGANISM: THE ASCIDIAN *Ciona intestinalis* ..... - 2 -

1.1 Tunicates: a peculiar branch in the tree of life ..... - 2 -

1.2 Ascidians embryonic development ..... - 3 -

1.3 *Ciona intestinalis* as model system in cellular and developmental biology ..... - 5 -

1.4 The larval ascidian nervous systems ..... - 6 -

1.4.1 The larval ascidian nervous systems in comparison with vertebrate CNS ..... - 9 -

1.5 A closer look to the ascidian sensory vesicle: the sensory organs ..... - 10 -

1.6 *Ciona* ocellus and its similarities with vertebrate eye ..... - 12 -

1.7 A marker of photoreceptor cell lineage in *Ciona*: what about Gsx? ..... - 15 -

GSX, A PARAHOX TRANSCRIPTION FACTOR ..... - 16 -

1.8 ParaHox genes in animal kingdom ..... - 16 -

1.9 *Gsx*, a gene as "ancient" as metazoans. Branch by branch in the animal phylogeny: an overview of the current knowledge. ..... - 19 -

1.9.1 Beyond the bilaterians: Gsx in the Cnidaria phylum ..... - 19 -

1.9.2 Gsx in protostomes: Lophotrochozoa ..... - 20 -

1.9.3 Gsx in protostomes: Ecdysozoa ..... - 21 -

1.9.4 Gsx in non-chordate deuterostomes: Ambulacraria ..... - 24 -

1.9.5 Gsx in chordate deuterostomes: vertebrates ..... - 24 -

1.9.6 Gsx in chordate deuterostomes: cephalochordates and urochordates ..... - 27 -

1.10 To sum up... ..... - 29 -

AIM OF THE THESIS ..... - 30 -

Chapter 2. Materials and Methods ..... - 31 -

2.1 *Ciona intestinalis* eggs and embryos collection ..... - 31 -

2.2 Chemical dechorionation and *in vitro* fertilization ..... - 31 -

2.3 Transgenesis *via* electroporation ..... - 31 -

2.4 Embryos observation and imaging analyses ..... - 32 -

2.5 PCR amplification from genomic or plasmid DNA ..... - 32 -

2.6 DNA gel electrophoresis ..... - 33 -

2.7 DNA gel extraction ..... - 33 -

2.8 DNA digestions with restriction endonucleases ..... - 34 -
2.9 Fill in of 5' overhangs ..... - 34 -
2.10 DNA dephosphorylation ..... - 34 -
2.11 DNA ligation ..... - 35 -
2.12 Bacterial cells electroporation ..... - 35 -
2.13 PCR screening ..... - 35 -
2.14 Plasmid DNA Mini- and Maxi-preparation ..... - 36 -
2.15 Sequencing ..... - 36 -
2.16 Oligonucleotides synthesis ..... - 37 -
2.17 Digested plasmids purification ..... - 37 -
2.18 Ribonucleic probes preparation ..... - 37 -
2.18.1 RNA labelling ..... - 37 -
2.18.2 Ribonucleic probes quantification by Dot Blot analysis ..... - 38 -
2.19 Whole Mount In situ Hybridization (WMISH) assays ..... - 39 -
2.19.1 Embryos preparation ..... - 39 -
2.19.2 WMISH protocol for single and double in situ ..... - 39 -
2.20 Preparation of constructs ..... - 42 -
2.21 *In silico* analysis of putative *trans*-acting factors ..... - 49 -
2.22 Site directed mutagenesis of putative Sox binding site on pGsxIII regulatory region ... - 50 -
2.23 Quantitative Real-time PCR (qPCR) ..... - 51 -
2.23.1 RNA extraction and RNA quality detection ..... - 51 -
2.23.2 cDNA synthesis ..... - 53 -
2.23.3 Quantitative Real-time PCR (qPCR) ..... - 53 -
2.24 Alignments of Gsx sequences ..... - 56 -
Chapter 3. Results ..... - 57 -
Ci-GSX PROTEIN SEQUENCE ..... **- 57 -**
3.1 Comparative analysis of Ci-gsx protein sequence ..... - 57 -
Ci-GSX EXPRESSION PATTERN ..... **- 61 -**
3.2 Refining Ci-gsx expression pattern ..... - 61 -
3.2.1 Confirming previous data: Ci-gsx expression pattern ..... - 61 -
3.2.2 Comparison of Gi-gsx and Ci-Six 3/6 expression pattern ..... - 62 -
3.2.3 Comparison of Gi-gsx, Ci-Meis and Ci-Ptf1a expression pattern ..... - 64 -
3.2.4 Comparison of Ci-gsx and Ci-Rx expression pattern ..... - 67 -
DEFINING Ci-GSX ROLE IN PHOTORECEPTOR CELLS PRECURSORS ..... **- 69 -**
3.3 Isolation of Ci-gsx regulatory region ..... - 69 -

3.4 Perturbing the activity of the endogenous Ci-gsx ..... - 72 -

3.4.1 Preliminary conclusions (Part I) ..... - 76 -

STUDIES ON Ci-GSX REGULATORY REGION ..... - 77 -

3.5 Dissection of *Ci-gsx* promoter ..... - 77 -

3.6 Searching for putative transcription factor binding sites in pGsxIII regulatory region (part I) ..... - 80 -

3.6.1 Ci-SoxB1 and Ci-SoxC expression pattern ..... - 81 -

3.6.2 Verifying potential Sox binding sites: further dissection of the promoter ..... - 82 -

3.7 Looking for the missing piece: fragmenting the fragment ..... - 86 -

3.8 Searching for putative transcription factor binding sites in pGsxIII regulatory region (part II) ..... - 87 -

3.8.1 Comparison between Ci-gsx and Ci-msxb expression pattern ..... - 88 -

3.8.2 Ectopic expression of Ci-msxb: effects on Ci-gsx ..... - 88 -

3.9 Preliminary conclusions (Part II) ..... - 90 -

Chapter 4. Discussion ..... - 91 -

4.1 *Ci-gsx* in the developing *Ciona intestinalis* Central Nervous System ..... - 91 -

4.2 A search for Ci-gsx function in Ciona intestinalis ..... - 102 -

4.2.1 Ci-gsx sequence analysis ..... - 102 -

4.2.2 Ci-gsx functional studies ..... - 104 -

4.3 Analysis of *Ci-gsx* regulatory region ..... - 109 -

4.3.1 Potential involvement of Sox transcription factors in Ci-gsx transcriptional activation ..... - 110 -

4.3.2 Potential involvement of Ci-Msxb in Ci-gsx transcriptional repression ..... - 116 -

4.4 Conclusions and future directions ..... - 118 -

REFERENCES ..... - 121 -

PREFACE

This thesis is the result of my PhD conducted from December 2010 to November 2013 at the Cell and Developmental Biology Laboratory at the Stazione Zoologica Anton Dohrn in Napoli, under the supervision of Dr Antonietta Spagnuolo. The thesis contains four chapters. Chapter 1 is a general introduction to the topic of the thesis, and includes a short overview on the model system, *Ciona intestinalis*, and on the current knowledge about Gsx transcription factor throughout the animal kingdom. Chapter 2 contains a detailed description of the materials and experimental procedures used during this study. Chapter 3 describes the results of the experiments, providing evidences of Ci-gsx function and transcriptional regulation during *Ciona* development. Chapter 4 discusses and summarizes the results obtained and also includes future perspectives.

LIST OF FIGURES

Chapter 1. Introduction

Figure 1.1 Phylogenetic relationships of deuterostomes....2

Figure 1.2 The ascidian *Ciona intestinalis*....3

Figure 1.3 Developmental fate restriction in ascidian embryos....4

Figure 1.4 Cell lineages of the ascidian larval CNS....8

Figure 1.5 Tripartite models of ascidian larval CNS....9

Figure 1.6 Sensory organs in *Ciona* sensory vesicle....11

Figure 1.7 Schematic comparison between the main building blocks of vertebrate eye and ascidian ocellus....13

Figure 1.8 Schematic representation of cell lineage of pigment sensory organs during *Ciona intestinalis* development....15

Figure 1.9 Evolution of the ParaHox cluster in animal phylogeny....18

Figure 1.10 Interactions between ind, msh and vnd factors....21

Figure 1.11 Interactions between Gsh2, Msx and Nk6.1 factors....26

Figure 1.12 Expression pattern of *Ci-gsx*....28

Chapter 2. Materials and Methods

Figure 2.1 Deletion of *pGsx* regulatory region....45

Figure 2.2 Preparation of pGsx>GsxHD>VP16 construct....47

Figure 2.3 Preparation of pGsx>Msx construct....49

Figure 2.4 Summary of the electrophoresis profiles generated by the BioAnalyzer....52

Chapter 3. Results

Figure 3.1 Clustal O aligment of protein sequence of Gsx family members....61

Figure 3.2 *Ci-gsx* expression pattern during *Ciona intestinalis* development....62

Figure 3.3 Positional relationship of *Ci-gsx* and *Ci-Six3/6* gene expression....63

Figure 3.4 Positional relationship of *Ci-gsx* and *Ci-Meis* gene expression.....65-66

Figure 3.5 *Ci-Rx* expression pattern.....67

Figure 3.6 Positional relationship of *Ci-gsx* and *Ci-Rx* gene expression.....68

Figure 3.7 Isolation of a cis-regulatory element required for *Ci-gsx* activation in the endogenous territories.....69

Figure 3.8 In vivo analysis of *pGsx>mCherry* construct.....71

Figure 3.9 Co-electroporation experiment of *pGsx>mCherry* and *pRx>GFP*.....70

Figure 3.10 WMISH hybridization with *Ci-Rx* probe on transgenic embryos electroporated with Ci-gsx interfering constructs.....73

Figure 3.11 WMISH hybridization with *Ci-Rx* probe on transgenic embryos electroporated with Ci-gsx interfering constructs.....73

Figure 3.12 Analysis of *pRx* activity after interfering with the endogenous *Ci-gsx* function.....74

Figure 3.13 Change in gene expression level assessed by qPCR on *pGsx>GsxHD>VP16* and *pGsx>GsxHD>WRPW* transgenic embryos.....75

Figure 3.14 Summarizing scheme of the results obtained from the *Ci-gsx* promoter region analysis.....79

Figure 3.15 *In silico* analysis of pGsxIII region.....80

Figure 3.16 *Ci-SoxB1* expression during *Ciona intestinalis* development.....81

Figure 3.17 Positional relationship of Ci-SoxB1 and Ci-gsx in early developmental stages.....82

Figure 3.18 *Ci-SoxC* expression as reported in ANISEED database.....82

Figure 3.19 Dissection of *Ci-gsx* regulatory region.....83

Figure 3.20 Scheme of the results obtained by site-directed mutagenesis of the three predicted Sox binding sites in pGsxIII region.....85

Figure 3.21 Dissection of the region included between pGsx13 and pGsx14 fragments.....86

Figure 3.22 Localization of the two putative Msx binding sites *in silico* (Genomatix software) identified in *pGsxIII* region.....87

Figure 3.23 Positional relationship of Ci-msxb and Ci-gsx.....88

Figure 3.24 Ectopic expression of *Ci-msxb in Ci-gsx* territories.....89

Chapter 4. Discussion

Figure 4.1 Cell lineages of Ciona CNS territories.....92-93

Figure 4.2 The expression of mCherry reporter in a transgenic larva electroporated with *pGsx>mChe* could include DA cells and group III photoreceptors.....96

Figure 4.3 *Ci-FoxB* expression pattern, reported in ANISEED database, compared to *Ci-gsx*.....98

Figure 4.4 Examples of *Gsx* expression pattern in bilaterians.....101

LIST OF TABLES

Chapter 2. Materials and Methods

Table 2.1 Genes of which ribonucleic probes have been synthesized.....38

Table 2.2 Oligonucleotides used for *pGsx* cloning.....43

Table 2.3 Oligonucleotides used for pGsx dissection.....44

Table 2.4 Oligonucleotides used for *pGsxIII* dissection.....46

Table 2.5 Oligonucleotides used for GsxHD amplification.....47

Table 2.6 Oligonucleotides used for Ci-msxb 5' amplification.....49

Table 2.7 Mutagenic oligonucleotides.....51

Table 2.8 Oligonucleotides used for qPCR.....55

Table 2.9 Accession Numbers of Gsx protein sequences used in the multiple alignment.....56

Chapter 1

INTRODUCTION

Life on our planet depends on the sun for energy, thus it is not surprising that the ability to detect and respond to light is the common feature of many living organisms, from single-cell algae and plants up to metazoans.

Light sensitive cells, the so-called photoreceptor cells, firstly appeared almost 600 million years ago and immediately became precious devices for survival advantages, like phototaxis and predator evasion by shadow detection. Over the following 100 million years these simple light-sensitive cells slowly and progressively improved their capacity to detect light, getting more sensitive and faster in light response. Then, with the Cambrian explosion (around 540 million years ago), the rapid evolution of animal body plan resulted into a great variety of visual systems that can be observed nowadays in nature (Lamb *et al.*, 2007). This high diversity led to the theory according to which eyes developed in all animal *phyla* 40 to 60 times independently (Salvini-Plawen and Mayr, 1961).

However, more recent data from genetic studies point to a monophyletic origin of the various eye types followed by divergent, parallel and convergent evolution (Gehring, 2005). As a matter of fact, beyond their diversity, the "visual systems" of different animals share some similarities at the functional and, more importantly, at the molecular level. The best known example in this context is probably represented by Pax6, an homeobox-containing transcription factor, considered the master control gene for eye morphogenesis and development throughout the animal kingdom (Gehring and Ikeo, 1999; Gehring, 2004). On these grounds, Gehring and Ikeo have suggested that the prototypic eyes, consisting of two cells, a photoreceptor and a pigment cell, were accidentally controlled by Pax6 and that the different types of eyes were then build up by recruiting, time by time in the course of evolution, new factors on this original genetic draft.

Whatever is the most appropriate theory, it is clear that the evolutionary histories of "the organs able to perceive light" and, in particular, the achievement of "an organ so wonderful as vertebrate camera eye" remains a fascinating and complex-to-unravel topic. A possible way to approach the question of eye evolution is to conduct a comparative analysis of the molecular

combinatorial codes responsible for the development of these structures along the animal kingdom, with a special attention to the less complex photosensitive structures present in lower invertebrate chordates of high comparative interest. In this context the ascidian *Ciona intestinalis,* a consolidate model system to address developmental as well as comparative-evolutionary questions, can be considered a valuable tool, as discussed in the following paragraphs.

THE MODEL ORGANISM: THE ASCIDIAN *Ciona intestinalis*

1.1 Tunicates: a peculiar branch in the tree of life

Ciona intestinalis belongs to the Tunicate (or Urochordate) *subphylum* that, together with Cephalochordate (amphioxus) and Vertebrate *subphyla*, constitute the Chordate *phylum*. Traditional studies, mainly based on morphological analyses, positioned cephalochordates as the closest living relatives of vertebrates and tunicates were placed in a basal evolutionary position (Schubert *et al.*, 2006; Beaster-Jones *et al.*, 2008). Recently, taking advantage of genomic sequences of many different species, including one cephalochordate and four tunicates, Delsuc and coworkers assembled a phylogenomic data set of 146 nuclear genes from 14 deuterostomes and 24 other slowly evolving species as an outgroup. The phylogenetic analyses of this data set provided compelling evidence that "Tunicates and not Cephalochordates are the closest living relatives of Vertebrates" (***figure 1.1***; Delsuc *et al.*, 2006; Vienne and Pontarotti, 2006; Dunn *et al.*, 2008; Putnam *et al.*, 2008).

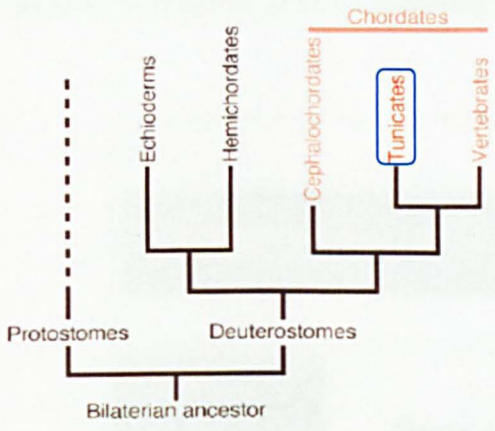


Figure 1.1 Phylogenetic relationships of deuterostomes. Ascidians belong to Tunicates *subphylum* (blue circle). (Adapted from Sasakura *et al.*, 2007).

The Tunicate *subphylum* is constituted by diverse groups of animals all having in common the cellulose-containing tunic that covers their body (Nakashima *et al.*, 2004) and that is responsible

for their nomenclature. They are traditionally divided into three classes: ascidians (sea squirts), thaliaceans (salps) and appendicularians (larvaceans).

Ciona intestinalis belongs to the ascidians, which are, in the first part of their life, planktonic swimming larvae (***figure 1.2 A***). After substrate settlement, larvae undergo a process of metamorphosis, becoming adult sessile filter feeder animals (***figure 1.2 B***). The adult form of ascidians could hardly justify their close phylogenetic relationship with vertebrates, since none of the distinctive features of the chordate body plan is clearly present. It is indeed in larval stage that the basic landmarks of the *phylum*, such as the notochord, the dorsal hollow nerve cord and segmented muscles (Holland *et al.*, 2004), are recognizable (***figure 1.2 A***, on the right).

Ciona larvae possess an extremely simplified structure. They develop very quickly (within 24 hours after fertilization, at 18°C) following a well-known developmental plan.

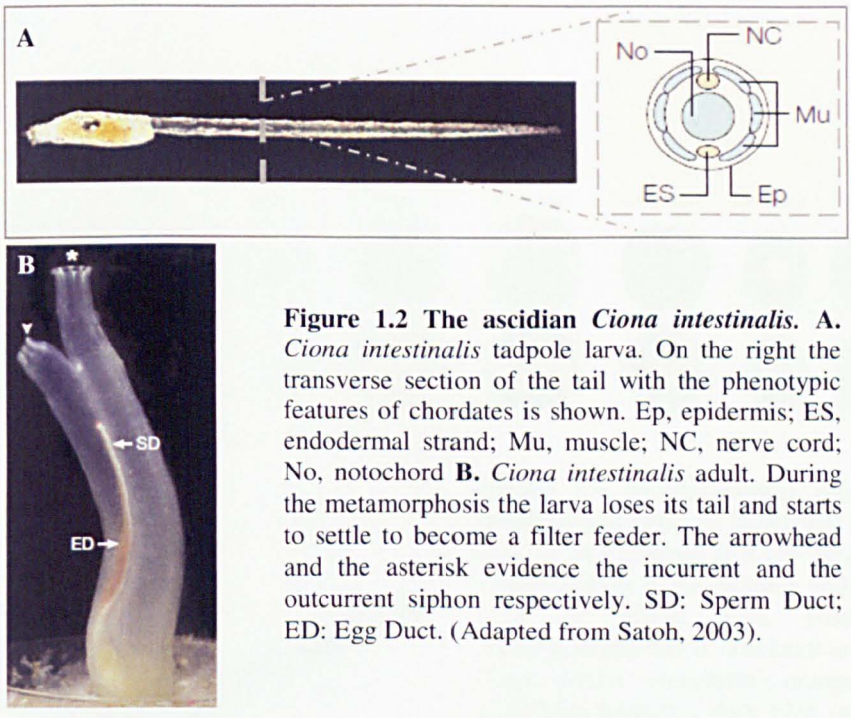


Figure 1.2 The ascidian *Ciona intestinalis*. A. *Ciona intestinalis* tadpole larva. On the right the transverse section of the tail with the phenotypic features of chordates is shown. Ep, epidermis; ES, endodermal strand; Mu, muscle; NC, nerve cord; No, notochord **B.** *Ciona intestinalis* adult. During the metamorphosis the larva loses its tail and starts to settle to become a filter feeder. The arrowhead and the asterisk evidence the incurrent and the outcurrent siphon respectively. SD: Sperm Duct; ED: Egg Duct. (Adapted from Satoh, 2003).

1.2 Ascidians embryonic development

Classically, ascidians embryogenesis has been considered a typical example of mosaic development, in which embryonic territories are specified from the zygotic stage throughout ooplasmic segregation of maternal determinants (Conklin, 1905a and Conklin 1905b). Blastomeres dissociated from an early cleavage stage, until the 110-cell stage, show that the

differentiation of epidermis, muscle and endoderm is autonomous without the involvement of any cell-cell signalling. In contrast, some ascidian territories, such as notochord, brain and pigment cells, failed to differentiate when embryonic cells were dissociated, suggesting that the development of those tissues is non-autonomous and requires cell interactions (Nishida, 1997). Ascidian embryonic development is very simple, fast and easy to follow. Soon after fertilization an invariant and bilaterally symmetrical cleavage program starts; the developmental stages of early ascidian embryos are named according to the number of cell, like 8-, 16-, 32-, 64-, 110-cell stages (***figure 1.3 A, B***; Satoh, 2003) and each blastomere is distinguishable with a specific and predictable lineage (Conklin, 1905c). The 8-cell stage embryo consists of the founder cells of four lineages, with the cells of the vegetal pole indicated by capital letters (A4.1 and B4.1) and animal cells named with small letters (a4.2 and b4.2; ***figure 1.3 A***; Conklin, 1905c). The embryo continues to cleave in a bilaterally symmetrical manner and thus each blastomere name refers to a pair.

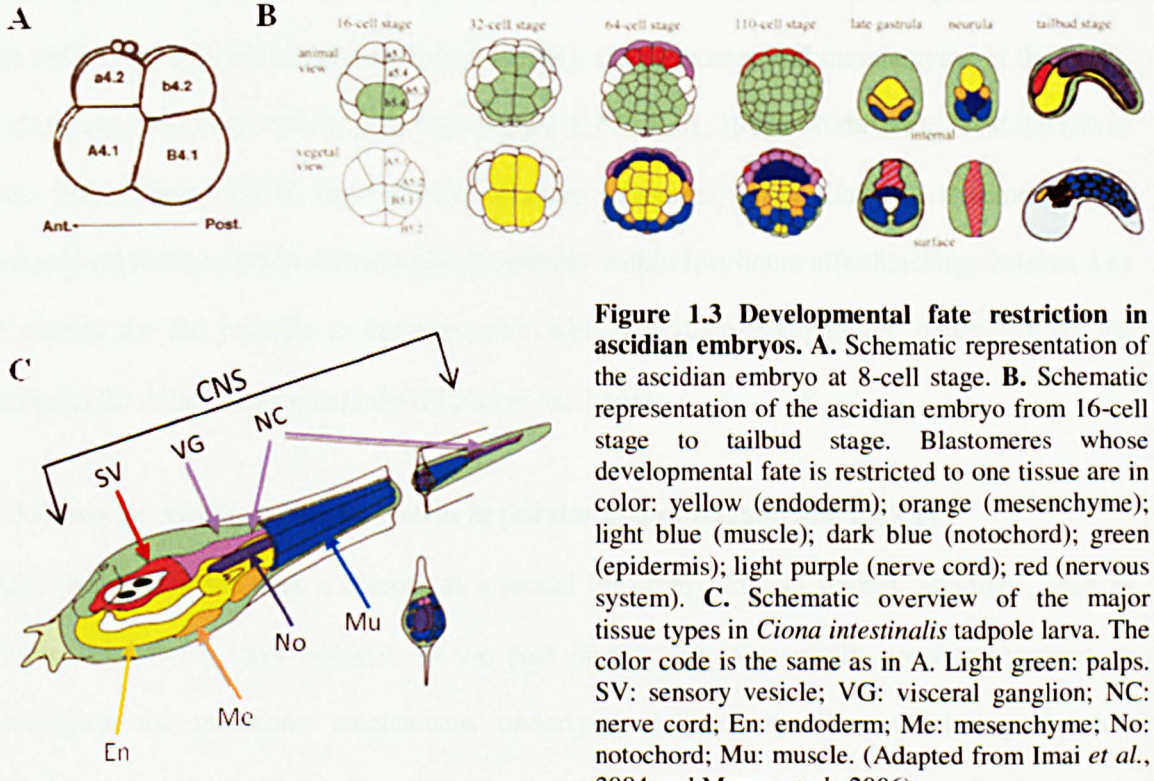


Figure 1.3 Developmental fate restriction in ascidian embryos. A. Schematic representation of the ascidian embryo at 8-cell stage. **B.** Schematic representation of the ascidian embryo from 16-cell stage to tailbud stage. Blastomeres whose developmental fate is restricted to one tissue are in color: yellow (endoderm); orange (mesenchyme); light blue (muscle); dark blue (notochord); green (epidermis); light purple (nerve cord); red (nervous system). **C.** Schematic overview of the major tissue types in *Ciona intestinalis* tadpole larva. The color code is the same as in A. Light green: palps. SV: sensory vesicle; VG: visceral ganglion; NC: nerve cord; En: endoderm; Me: mesenchyme; No: notochord; Mu: muscle. (Adapted from Imai *et al.*, 2004 and Munro *et al.*, 2006).

In contrast to the blastomeres of vertebrates, those of most ascidians are fate restricted early during development, from the beginning of gastrulation, shortly before the 110-cell stage (Munro *et al.*, 2006). At the tailbud stage the main tissues and organs that constitute the future larva begin to be differentiated (***figure 1.3 B***; Satou *et al.*, 2001).

As previously mentioned, *Ciona* larvae develop in a relatively short time, within 24 hours after fertilization. The formation of the tadpole larva involves all the morphogenetic movements responsible for the shaping of vertebrate embryos (*e.g.* convergent extension, invagination, cell migration, oriented cell divisions), nevertheless occurring within a little number of cells. Interestingly, cleavage patterns, cell lineages and the final body plan are remarkably conserved between distantly related ascidian genera, such as *Ciona* and *Halocynthia*, which diverged several hundred years ago (Hudson and Yasuo, 2008; Lemaire, 2009). Since the overall ontogeny of these organisms is very similar, cell lineage data obtained in one species can be often considered valid for another one.

Ciona larvae are made of only ~2600 cells, organized into a small number of organs that include the epidermis, the central nervous system (CNS), the endoderm and mesenchyme in the trunk, and the notochord and muscle in the tail (***figure 1.3 C***; Katz, 1983). At the rostral end the larvae bear palps (***figure 1.3 C***, light green), adhesive structures which allow the attachment to a suitable substrate to begin metamorphosis, usually within few hours after hatching. It takes 2 or 3 months for the juvenile to become adult with reproductive capability, depending on the temperature of the environment (Marikawa *et al.*, 1994).

1.3 *Ciona intestinalis* as model system in cellular and developmental biology

After being used for over a century as a model for embryological studies, ascidians, such as *Ciona intestinalis*, have become, in the past decade, an increasingly popular organism to investigate the molecular mechanisms underlying cell-fate specification during chordate development. *Ciona* embryos indeed permit detailed visualization of gene expression by whole-mount *in situ* hybridization (Satoh, 2001). The clonal restriction of larval tissues occurs early,

within 64-110 cell stage, and the lineage leading to the formation of larval tissues has been depicted almost completely (Nishida, 1987). *Ciona* genome is quite small and compact (160 Mb, 1/17 of human genome), being composed by about 16000 protein coding sequences, with very few duplicated genes (Dehal *et al.*, 2002).

In addition, a large quantity of cDNA and EST (Expressed Sequence Tag) information substantiates the gene models and packages of cDNA clones, named "*Ciona* gene collection", are freely available in the tunicate community. An important characteristic of *Ciona* genome is that most of the promoters are relatively short and usually located in close proximity to the transcription start sites of the genes (Corbo *et al.*, 1997; Takahashi *et al.*, 1999; Fanelli *et al.* 2003; Alfano *et al.*, 2007; Squarzoni *et al.*, 2011). Furthermore, comparisons between complementary *C. intestinalis* and *C. savignyi* sequences have indicated that these two species are at sufficient evolutionary distance to permit efficient identification of conserved regulatory sequence information (Bertrand et al., 2003; Johnson et al., 2004; Squarzoni et al., 2011). These peculiarities, combined with the method of transgenesis by electroporation that permits the transformation of hundreds or even thousands of embryos simultaneously (Corbo *et al.*, 1997), have made *Ciona* particularly useful for studies of transcriptional regulation. This technique allows also to create "knock-out" phenotypes, by expressing dominant negative forms of particular genes using identified lineage-specific enhancers. Thus, all these propitious features, coupled with a simplified chordate body plan that contains rudiments of most vertebrate tissues, make *Ciona* a very suitable model organism to explore the genetic circuitry responsible for the establishment of typical basic chordate features, such as the development and compartmentalization of the nervous system.

1.4 The larval ascidian nervous systems

The larval nervous system of *Ciona* is simple but well differentiated, mainly characterized by considerable similarities with the vertebrate CNS. Of particular interest is the dorsal tubular nervous system in *Ciona* embryo that derives from the rolling of a flat neural plate into a hollow neural tube (Nishida, 1986; Nicol and Meinertzhagen, 1988a) in a manner resembling

amphibian embryos and thus closely conserved during vertebrates neurogenesis (Cole and Meinertzhagen, 2004).

Ciona CNS contains fewer than 400 cells, about one-third of which accounted as neurons and the remaining cells assumed to be non-neuronal glial cells (Nicol and Meinertzhagen, 1991; Imai and Meinertzhagen, 2007). It is subdivided, along the longitudinal axis, in three main regions: the anterior sensory vesicle, connected by the neck to the visceral ganglion, and the dorsal nerve cord posteriorly (***figure 1.3 C***, red and purple; Katz, 1983; Nicole and Meinertzhagen, 1991).

The sensory vesicle is a hollow cavity sited in the trunk of the larva. As later described in detail, it contains different sensory receptor cells, mostly localized in two pigmented organs, the anterior otolith and the more posterior ocellus (Dilly, 1969a; Eakin and Kuda, 1971; Ohtsuki, 1991; Tsuda et al., 2003a), which guide the swimming behavior of the tadpole larvae (Tsuda et al., 2003b).

The neck, composed by only six cells, whose function is still unclear, links the sensory vesicle to the visceral ganglion. The latter probably works as the integrative center of the nervous system, receiving inputs from the various sensory structures located anteriorly. It includes 45 cells with 18 neurons, of which five pairs are motoneurons. The tubular caudal nerve cord extends along the entire larval tail; it is formed by four longitudinal rows of ependymal cells (dorsal, ventral and two lateral), enclosing a neural canal (Nicol and Meinertzhagen, 1991; Meinertzhagen and Okamura, 2001).

The cell lineages of the ascidian CNS during embryonic development have been quite well described (Nishida, 1987; Nicol and Meinertzhagen, 1988a; Nicol and Meinertzhagen, 1988b; Cole and Meinertzhagen, 2004; Taniguchi and Nishida, 2004). It derives from the a-, b- and A-lineages, which contribute to different parts of the CNS (***figure 1.4***).

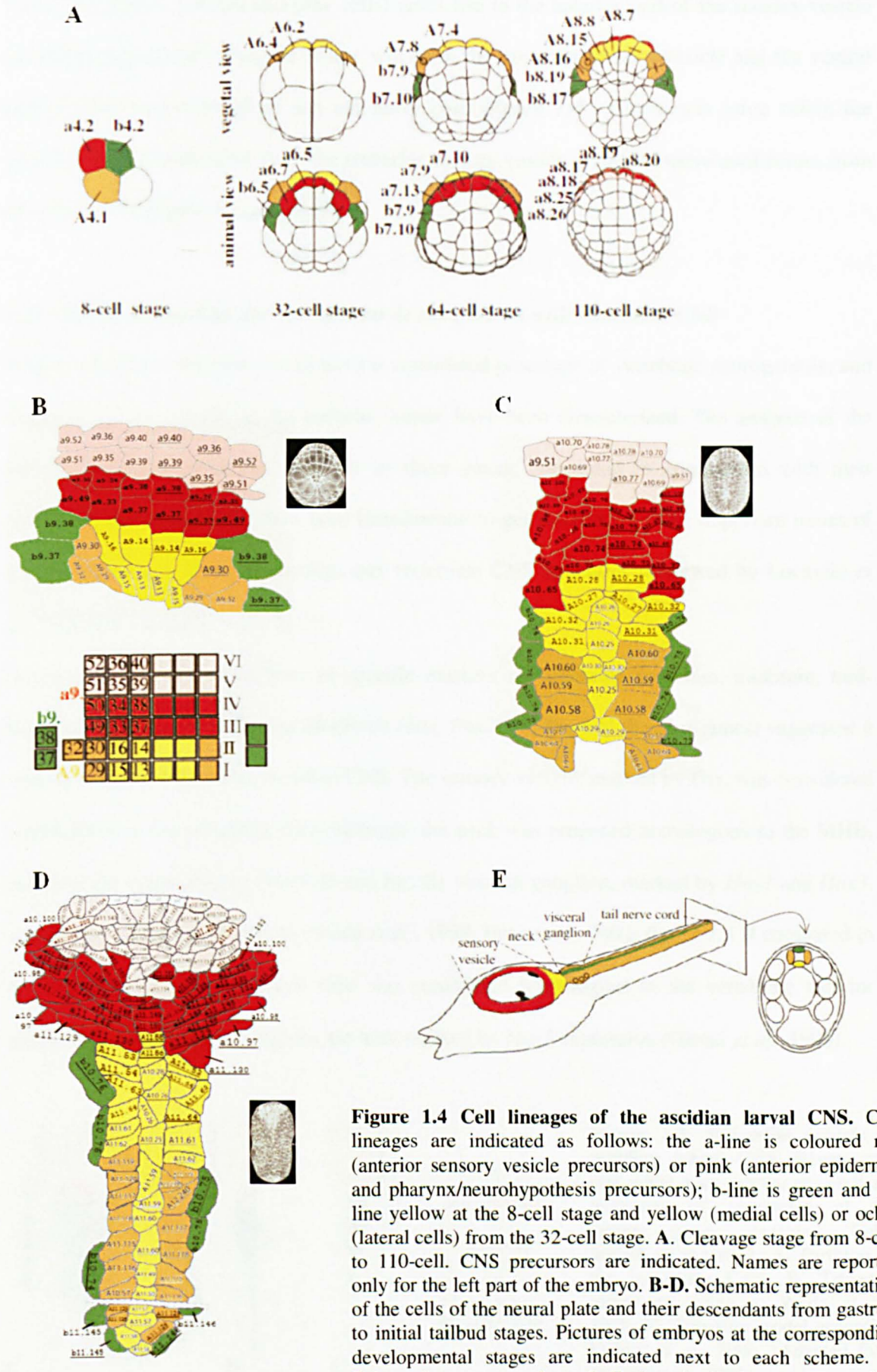


Figure 1.4 Cell lineages of the ascidian larval CNS. Cell lineages are indicated as follows: the a-line is coloured red (anterior sensory vesicle precursors) or pink (anterior epidermis and pharynx/neurohypothesis precursors); b-line is green and A-line yellow at the 8-cell stage and yellow (medial cells) or ochre (lateral cells) from the 32-cell stage. **A.** Cleavage stage from 8-cell to 110-cell. CNS precursors are indicated. Names are reported only for the left part of the embryo. **B-D.** Schematic representation of the cells of the neural plate and their descendants from gastrula to initial tailbud stages. Pictures of embryos at the corresponding developmental stages are indicated next to each scheme. **E.** Schematic representation of *C. intestinalis* larvae. The colors indicate the different embryonic origin of each part of the CNS. (Adapted from ANISEED database and from Hudson and Yasuo, 2005).

The a-line (***figure 1.4***, red and pink cells) gives rise to the anterior part of the sensory vesicle (including pigmented cells); the A-line will form the posterior sensory vesicle and the ventral parts of the visceral ganglion and tail nerve cord (***figure 1.4*** , yellow and ochre cells); the dorsal-most cell of the CNS from the posterior sensory vesicle to the tail nerve cord comes from the b-line cells (***figure 1.4***, green cells).

1.4.1 The larval ascidian nervous systems in comparison with vertebrate CNS

With its chordate affinities, *Ciona* larva is considered prototype of vertebrate neurogenesis, and numerous genes specific to the nervous system have been characterized. The analysis of the expression pattern territories of some of these genes, conducted in comparison with their vertebrate counterparts, has then been instrumental to get insights on some important issues of regional homology between ascidian and vertebrate CNS structures (reviewed by Locascio *et al.*, 2009 and Sasakura *et al.*, 2012).

Initially, the expression analysis of specific markers of vertebrate forebrain, midbrain, mid-hindbrain boundary (MHB) and hindbrain (*Otx*, *Pax2/5/8*, *En*, and the *Hox* genes) suggested a tripartite organization of the ascidian CNS. The sensory vesicle, marked by *Otx*, was considered homologous to the vertebrate fore-midbrain; the neck was proposed homologous to the MHB, based on the expression of *Pax2/5/8* and *En*; the visceral ganglion, marked by *Hox1* and *Hox3*, was assimilated to the hindbrain (Wada *et al.*, 1998; Imai *et al.*, 2002; ***figure 1.5 B*** compared to ***A***). The ascidian anterior neural tube was considered homologous to the vertebrate anterior spinal cord since these two regions are both marked by *Hox5* expression (Gionti *et al.*, 1998).

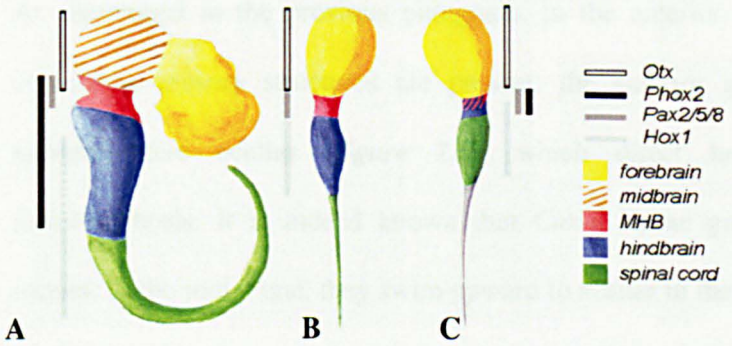


Figure 1.5. Tripartite models of ascidian larval CNS. Black, grey and withe bars graphically indicate territories of expression of the marker genes elencated. Color legend on the right. **A.** Partition of the vertebrate CNS. **B.** Tripartite model according to Wada *et al.*, 1998. **C.** Tripartite model according Dufour *et al.*, 2006. (Adapted from Dufour *et al.*, 2006).

Later, Dufour *et al.* (2006) re-examined this tripartite organization of *Ciona* CNS by using the homeobox gene *Phox2* as an additional marker. They suggested that the neck region of *Ciona* corresponds non only to the MHB, as previously mentioned, but to MHB plus hindbrain of vertebrates, because the posterior part of the neck expresses *Pax2/5/8*, *Hox1* and *Phox2*, the same gene expression profile as the hindbrain (***figure 1.5 C*** compared to ***A***). In this model the visceral ganglion of *Ciona* was then assimilated to the most anterior part of the spinal cord, which is marked by *Hox1*.

Thought that future and further analyses will undoubtedly permit to define much better the right orthology between ascidian and vertebrate CNS, it is noteworthy that a recent detailed characterization of gene networks in the developing CNS strongly supports the presence of the MHB organizer already in ascidians. The MHB is a structure that has been for long regarded as a novel property of the vertebrate CNS and considered to be mostly involved in the elaboration of the more complex vertebrate brain (Liu and Joyner, 2001; Wurst and Bally-Cuif, 2001; Rhinn *et al.*, 2006). In the vertebrate CNS the Fgf8 signalling molecule is expressed in the MHB organizer and, by controlling a network of factors, plays a key role in delineating separate midbrain and hindbrain compartments. The authors demonstrated that in *Ciona* embryos the Fgf8 ortholog, FGF 8/17/18, defines the posterior sensory vesicle and the neck region by generating regional patterns of gene expression (Imai *et al.*, 2009), thus suggesting that the primitive MHB-like activity predates the vertebrate CNS.

1.5 A closer look to the ascidian sensory vesicle: the sensory organs

As mentioned in the previous paragraph, in the anterior sensory vesicle two pigment cell-containing sensory structures are present: the anterior geotactic otolith and the posterior photoreceptive ocellus (***figure 1.6***), which direct larval swimming behavior before metamorphosis. It is indeed known that *Ciona* larvae go through two different swimming phases: in the initial one, they swim upward to scatter in the sea; in the second phase they swim or sink downwards to settle and metamorphose (Tsuda *et al.*, 2003b).

The otolith and ocellus organs can easily be distinguished based on their morphology and their specific type of pigmentation (Nishida and Satoh, 1989).

The otolith (or statocyte) is a single-cell organ, lying in the anterior part of the sensory vesicle, simply composed by a large spherical pigment granule within a vacuole, measuring about 10-15 µm in diameter (Sakurai *et al.*, 2004) and connected by a narrow stalk to the ventral floor of the sensory vesicle. It is involved in gravity perception, controlling the upward swimming behavior of the larvae thanks to the movement of the melanin granule, which supplies gravity information to the animal (Ohtsuki, 1991; Tsuda *et al.*, 2003b).

The ocellus, located in the right posterior wall of the sensory vesicle, is responsible for the negatively phototactic, downward swimming behavior of *Ciona* larvae. It shows a higher degree of structural complexity, compared to the otolith, with its multicellular organization composed by a pigment cell, three lens cells and about 30 photoreceptor cells (PRCs; ***figure 1.6 C***).

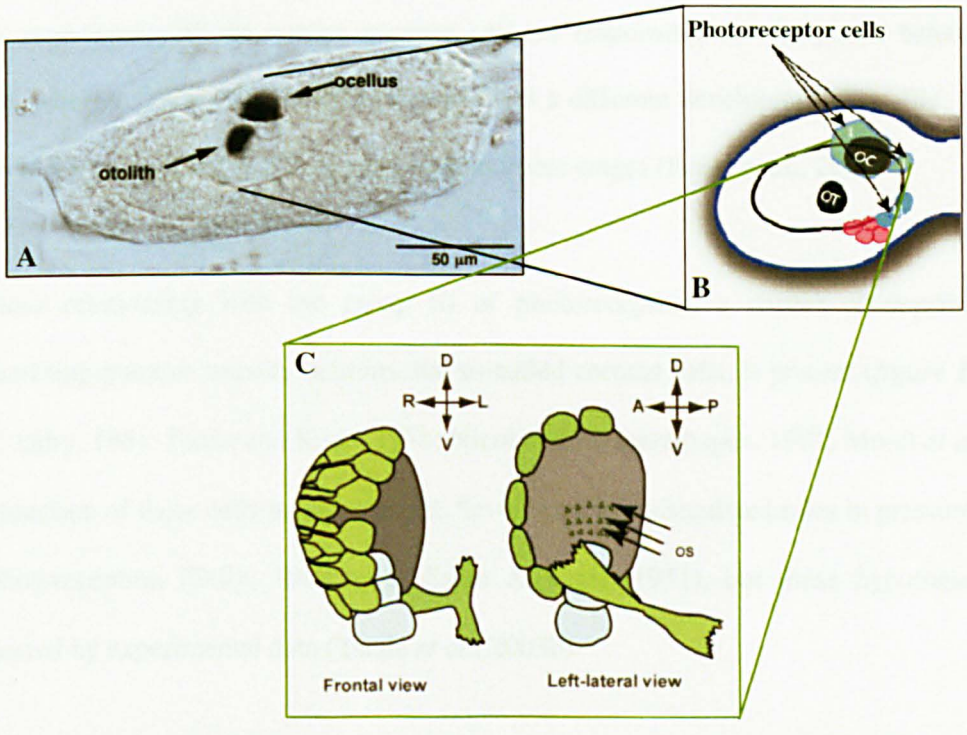


Figure 1.6 Sensory organs in *Ciona* sensory vesicle. A. Larval trunk. Otolith and ocellus are indicated by black arrows (Tsuda *et al.*, 2003b). **B.** Schematic representation of larval sensory vesicle. The pigment cells of Otolith (OT) and Ocellus (OC) are grossly shown in black. The light blue spots indicate the three groups of photoreceptor cells (PRCs) present in the sensory vesicle. The coronet cells, in close relationship with the population III of PRCs, are indicated in pink (Adapted from Razy-Krajka *et al.*, 2012). **C.** Graphical scheme of the frontal view and lateral view of the ocellus. Photoreceptor cells are depicted in green, pigment cell in black and lens cells in light blue. Os: outer segments. (Horie *et al.*, 2005).

The ocellus pigment cell is cup shaped and contains several membrane-bound pigment granules, which exert a photoprotective role for the posterior photoreceptor cells, filtering directionally the incoming light (Tsuda et al., 2003b). Photoreceptor cells are sited in the right, posterior wall of the sensory vesicle (Horie *et al.*, 2005), while the three lens cells are located in the bottom part of the ocellus structure, between photoreceptors and pigment cell (***figure 1.6 C***).

A relatively recent study (Horie *et al.*, 2008) demonstrated that the photoreceptor cells associated to the pigmented ocellus can be divided in two morphological distinct groups of cells, named group I and II, according to the location of the outer segments inside or outside the pigment cup, respectively. Furthermore, a group III of photoreceptor cells in the left ventral part of the brain vesicle, in proximity to the otolith and apart from the ocellus pigment cell, has been identified (***figure 1.6 B***) and defined as a novel ocellus, lacking a pigment cell. Structural analyses coupled with functional studies, by laser ablation experiments, revealed that only the PRCs associated with the ocellus pigment cell are responsible for the photic behavior of the larvae, whereas the group III, which also exhibits a different developmental profile, could have a role in later larval stages and/or early metamorphic stages (Horie *et al.*, 2008).

In close relationship with the group III of photoreceptors, a cluster of dopamine (DA)-synthesizing putative sensory neurons, the so-called coronet cells, is present (***figure 1.6 B***, pink cells. Dilly, 1961; Eakin and Kuda, 1971; Nicol and Meinertzhagen, 1991; Moret *et al.*, 2005a). The function of these cells is not clear yet. Several authors speculated roles in pressure detection or photoreception (Dilly, 1969 a, b; Eakin & Kuda, 1971), but these hypotheses are not supported by experimental data (Tsuda *et al.*, 2003b).

1.6 *Ciona* ocellus and its similarities with vertebrate eye

The vertebrate eye is a complex optical system which collects light from the surrounding environment, focuses it through an adjustable assembly of lenses to form an image, converts this image into a set of electrical signals, that are transmitted to the brain through intricate

neural pathways. Despite its complex function and structure, the main building blocks of vertebrate eye are two: the neural retina and the RPE (Retinal Pigmented Epithelium; ***figure 1.7 A***). The retina contains the photoreceptor cells devoted to perceive light and convert it into signals; the RPE is a monolayer of pigmented cells that closely interacts with photoreceptors and, among a number of roles, is involved in the maintenance of visual function and in protecting photoreceptors from excess incoming light.

In comparison with vertebrate eye, the simplest "eye", such as *Ciona* ocellus, is only able to detect whether the surroundings are light or dark; nevertheless, its building blocks are two, as in vertebrates: the photoreceptor cells and the dark shielding pigment cell (***figure 1.7 B***). In addition, some aspects of photo-transduction and developmental mechanisms of *Ciona* ocellus appear to be shared with vertebrate eyes, thus indicating that these two structures could have been derived from a common archetypal "visual organ" (Kusakabe *et al.*, 2001; Sato and Yamamoto, 2001; Lamb *et al.*, 2007).

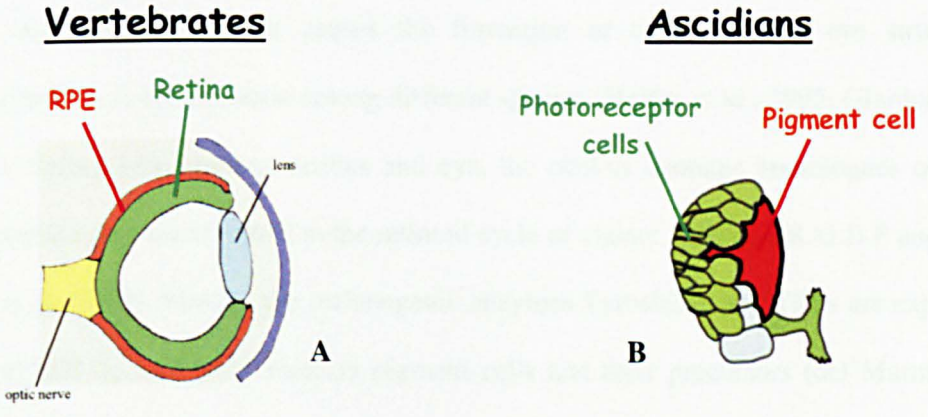


Figure 1.7 Schematic comparison between the main building blocks of vertebrate eye (A) and ascidian ocellus (B). Retinal Pigmented Epithelium of the eye (A) and pigment cell of the ocellus (B) are depicted in red. Retina of vertebrate eye (A) and photoreceptor cells of the ocellus (B) are depicted in green.

In particular, the light-response of *Ciona* photoreceptors is hyperpolarizing, as for vertebrates (Gorman *et al.*, 1971). Furthermore, the photo-transduction process in vertebrates requires the visual opsins, which are G-protein coupled receptors, and visual arrestins, small proteins needed to regulate opsin signal transduction (Arshavsky, 2002; Blomhoff and Blomhoff, 2006). Similarly, photo-transduction in *Ciona* uses opsins, precisely Ci-opsin1 (three opsin homologs

are present in *Ciona* genome) and Ci-arrestin (one arrestin homolog is present in *Ciona* genome), both expressed in ocellus photoreceptor cells as well as in the group III of photoreceptor cells (Kusakabe *et al.*, 2001; Nakagawa *et al.*, 2002; Nakashima *et al.*, 2003; Horie et al., 2008).

At the developmental level, it is known that in vertebrates Pax6 and RX are involved in eye formation. RX is indeed required for the eye field specification, since its targeted elimination causes total absence of the eye in mice (Bailey *et al.*, 2004). Likewise, in *Ciona* it has been reported that the homologous gene, *Ci-Rx*, is fundamental for ocellus development and function (D'Aniello *et al.*, 2006). Moreover, mutations in vertebrate and invertebrate *Pax6* gene result in defects or absence of the eye (Callaerts *et al.*, 1997). In tunicates *Pax6* is expressed in the nervous system territories, including pigment organs precursors, but its role has not been clarified yet (Mazet *et al.*, 2003). However, it has been demonstrated that the ectopic expression in *Drosophila* imaginal discs of *Pax6* homologous genes from either the ascidian *Phallusia mamillata*, mouse or *Drosophila* causes the formation of supernumerary eye structures, indicating a functional conservation among different species (Halder *et al.*, 1995; Glardon *et al.* 1997). As a further link between ocellus and eye, the ocellus contains homologues of three vertebrate proteins that are involved in the retinoid cycle of vision: RPE65, CRALB P and BCO (Takimoto *et al.*, 2006). Finally, the melanogenic enzymes Tyrosinase and TRPs are expressed in vertebrate RPE as well as in tunicate pigment cells and their precursors (del Marmol and Beermann, 1996; Caracciolo *et al.*, 1997; Sato *et al.*, 1997; Esposito *et al.*, 2012).

Collectively, all these data strongly support *Ciona* as a valuable model system to study, in a relatively simple way, the "basic" mechanism underlying the development of complex structures, such as the eyes, much more difficult to be approached in vertebrates.

1.7 A marker of photoreceptor cell lineage in *Ciona*: what about Gsx?

As previously mentioned, in *Ciona* extensive information are available on the cell lineage of most of the main tissues and organs, so that it is possible to follow many developmental processes. Concerning the lineage of photoreceptor cells, a detailed analysis of the mitotic history of CNS precursor cells led to infer that, since neural plate stage, the right a9.33 and a9.37 cells most probably represent the photoreceptor cells progenitors (Nicol and Meinertzhagen, 1991; Cole and Meinertzhagen, 2004; ***figure 1.8 A***, green cells). These cells, as schematized in the ***figure 1.8 A***, are localized just between the pigment cell precursor blastomeres, a9.49 pairs (blue cells in the figure). During neurulation the a9.49 couple go through one mitotic division, leading to the formation of a10.97 and a10.98 cell pairs (***figure 1.8 B***, blue cells), that converge and intercalate along the anterior-posterior axis during the neural tube closure. The couple a10.97 form the pigment cells of otolith (the anterior one) and ocellus (the posterior one) sensory organs, whereas the a10.98 cells form part of the sensory vesicle (***figure 1.8 C, D***, blue cells). Differently from the pigment cell precursors, the a9.33 and a9.37 blastomeres undergo several mitotic division, giving rise to almost 30 cells that, at larval stage, come close to the a10.97 posterior pigment cell to form the photoreceptors of the ocellus sensory organ (***figure 1.8 D***).

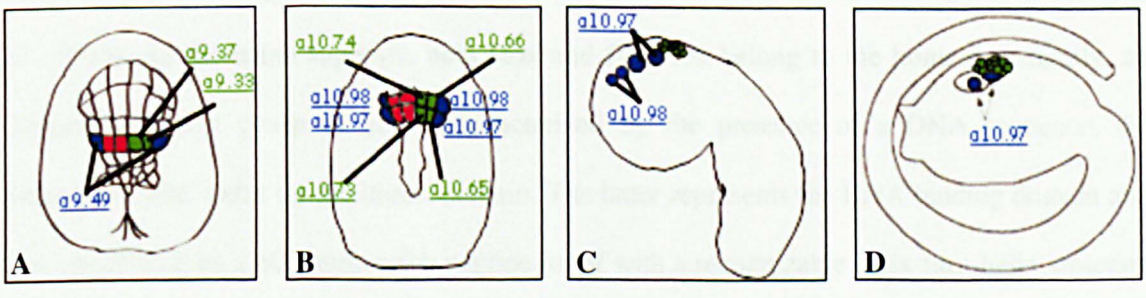


Figure 1.8 Schematic representation of cell lineage of pigment sensory organs during *Ciona intestinalis* development. Pigment cells precursors are depicted in blue; photoreceptor cells precursors are depicted in green; cells proposed to be coronet cells precursors are depicted in pink. **A.** Gastrula stage, dorsal view. **B.** Neurula stage, dorsal view. **C.** Tailbud stage, lateral view. **D.** Larval stage, lateral view.

In the same study Cole and Meinertzhagen (2004) inferred that some of the progeny from left a9.33 and a9.37 blastomeres (***figure 1.7 A, B***, pink cells) likely forms the almost 20 coronet

cells, which lie in proximity to the otolith and close to the group III of photoreceptors (Cole and Meinertzhagen, 2004).

At molecular level, besides the demonstration of a pivotal role played by *Ci-Rx* in the ocellus photoreceptor cells differentiation, from the early tailbud stage (D'Aniello *et al.*, 2006), very little information are available on the molecular pathways and factors responsible for early photoreceptor specification in *Ciona* and, more generally, in ascidians.

One potential strategy to tackle this issue would be querying available databases in order to identify genes "developmentally relevant" and potentially involved in the process, based on their expression profiles. A search on ANISEED database (http://www.aniseed.cnrs.fr/), which provides high throughput data on *Ciona intestinalis* and *in situ* experiment data from literature with more than 30.000 in situ hybridization profiles, led to the identification of the ParaHox gene Ci-*gsx* as a very interesting candidate, since it is expressed in the right place and at the right developmental time to participate in early photoreceptor differentiation.

GSX, A PARAHOX TRANSCRIPTION FACTOR

1.8 ParaHox genes in animal kingdom

Gsx, together with Xlox and Cdx, belongs to the ParaHox family of transcription factors, considered the sister group (paralogue) of the more extensively studied Hox genes (Brooke *et al.*, 1998). As the name suggests, both Hox and ParaHox belong to the homeobox family, an ancient and huge group of genes characterized by the presence of a DNA sequence, the homeobox, that codes for the homeodomain. The latter represents the DNA binding domain and it is constituted by a 60 amino acids peptide motif with a recognizable helix-turn-helix structure (McGinnis *et al.*, 1984; Scott and Weiner, 1984; Lewin, 2000). Most homeodomain proteins bind to short DNA sequences of only 6 bp, often with a common TAAT core followed immediately by two bases that confer specificity (Treisman *et al.*, 1989, 1992).

Hox and ParaHox genes probably evolved from a primitive ProtoHox cluster, likely duplicated in the last common ancestor of all animals, before the origin of *Porifera* (***figure 1.8***; Garstang and Ferrier, 2013; Mendivil Ramos *et al.*, 2012).

Much is known about the Hox genes, which fascinated biologists since their discovery in 1984 (Wakimoto *et al.*, 1984), because of their powerful and fundamental functions during embryonic development throughout the animal kingdom, providing the basis for anterior-posterior axis specification. Key typical traits of the Hox genes, at least partially conserved in ParaHox genes, are: the cluster organization in the genome; the spatial colinearity (the position of the genes within the cluster is related to their relative domains of expression along the major body axis), and/or the temporal colinearity (genes can be activated progressively during development according to their relative position in the cluster).

The identification of the ParaHox cluster, instead, is more recent and for this reason in many cases less data are available. "Only" in the 1998 indeed, *Gsx* (*genomic screened homeobox*), *Xlox* (*Xenopus laevis homeobox 8*) and *Cdx* (*Caudal-type homeobox*) were for the first time identified as a genomic cluster in the amphioxus *Branchiostoma floridae* (Brooke *et al.*, 1998). In amphioxus, ParaHox genes exhibits both temporal and spatial colinearity: *Gsx* is expressed in the anterior neural tube, *Xlox* and *Cdx* in the middle and posterior endoderm and neural tube respectively, with *Cdx* expressed first and *Gsx* last in terms of timing of gene activation (Brooke *et al.*, 1998; Ferrier *et al.*, 2005; Osborne *et al.*, 2009).

To date, it has been interestingly found that ParaHox cluster organization, among deuterostomes, is conserved in *Xenopus*, mouse, human, in the sea star *Patiria miniata* and the hemichordate *Ptychodera flava* (Ferrier *et al.*, 2005; Mulley, 2008; Annunziata *et al.*, 2013; Furlong and Ikuta *et al.*, 2013), but is lost in teleost fishes, in the ascidian *Ciona intestinalis* (Ferrier and Holland, 2002; Mulley *et al.*, 2006) and in the sea urchin *Strongylocentrotus purpuratus* (Arnone *et al.*, 2006). Cluster organization seems absent also in protostomes, such as the annelid *Platynereis dumerilii* (Hui *et al.*, 2009) and the insect *Drosophila melanogaster* (***figure 1.9***. Macdonald and Struhl, 1986; Weiss *et al.*, 1998).

Comparative studies, among different bilaterian *taxa*, confirmed the colinear expression of the three ParaHox genes along the anterior posterior axis of the developing CNS and gut. *Gsx* is

generally the anterior-most one; *Xlox* is expressed in the central regions of developing gut (such as pancreas of vertebrates) and CNS; *Cdx* is present in more posterior regions of CNS and gut. Exceptions, perhaps as a consequence of cluster break, are represented by *C. intestinalis* and *S. purpuratus*, in which the spatial colinearity is retained, but the temporal colinearity is lost in *Ciona* and reversed in the sea urchin (data reviewed by Annunziata *et al.*, 2013 and Ikuta *et al.*, 2013; ***figure 1.9***).

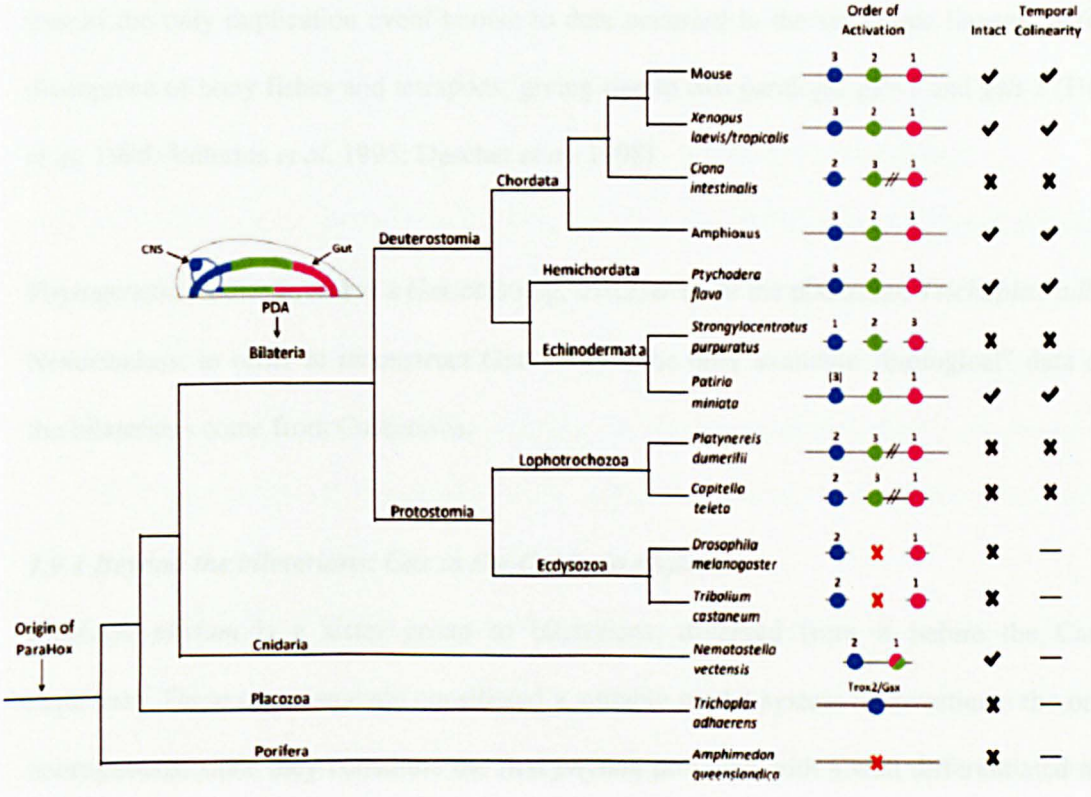


Figure 1.9 Evolution of the ParaHox cluster in animal phylogeny. Genomic organization of ParaHox genes for each species is shown. *Gsx* is depicted in blue, *Xlox* in green, *Cdx* in pink. The putative expression domains in the CNS and gut of the protostome-deuterostome (PDA) ancestor are schematized in the upper left of the tree. On the left side gene linkage is represented by a continuous line connecting genes. Double diagonals represent genes located on the same chromosome but separated by large distances. Red crosses indicate gene loss. The order in which ParaHox genes are activated and expressed has been indicated numerically (*P. miniata* Gsx activation is in parentheses due to presumed later larval expression).
The presence of an intact cluster or temporal colinearity is indicated by a black check or cross. Horizontal lines, when present, indicate that temporal colinearity cannot be resolved due to the absence of one or more ParaHox genes. (Adapted from Garstang and Ferrier, 2013).

1.9 *Gsx*, a gene as "ancient" as metazoans. Branch by branch in the animal phylogeny: an overview of the current knowledge.

Among the Hox/ParaHox gene families, *Gsx* is one of the most conserved throughout the animal kingdom, with orthologs found from placozoans up to vertebrates. Studies devoted to the reconstruction of its ancestral function and later diversification during evolution are facilitated by the relatively conservative history of *Gsx*, without intensive duplication events that should have easily resulted in functional diversification of the duplicated copies (Finnerty *et al.*, 2003). Indeed the only duplication event known to date occurred in the vertebrate lineage, before the divergence of bony fishes and tetrapods, giving rise to two paralogs, *gsh-1* and *gsh-2* (Hsieh-Li *et al.* 1995; Valerius *et al.* 1995; Deschet *et al.* 1998).

Phylogenetic studies revealed a *Gsx* ortholog, *trox2*, even in the placozoan *Trichoplax adherens*. Nevertheless, in order to reconstruct *Gsx* "story", the only available "biological" data outside the bilaterians come from Cnidarians.

1.9.1 Beyond the bilaterians: Gsx in the Cnidaria phylum

Cnidaria phylum is a sister group to bilaterians, diverged from it before the Cambrian explosion. These organisms are considered a suitable model system to investigate the origin of neurogenesis, since they constitute the first *phylum* provided with a well differentiated nervous system. Here *Gsx* ortholog, called *cnox2* (or *anthox2*), is the best characterized homeobox gene of the *phylum* (Finnerty *et al.*, 2003). Expression profile studies have been performed in numerous species, but a clear-cut consensus in the resulting pattern cannot be easily found, given the strong variation in *cnox2* location in each case. However, what seems to become more and more evident, thanks both to expression and functional studies, is that the ancestral gsx/cnox2 function should be related to neurogenesis and oral patterning. Referring to this, in *Hydra* it has been clearly demonstrated that cnox2, marking specifically the nervous system, promotes apical neurogenesis and head patterning during head regeneration after amputation. In particular, *cnox2* specifically identifies a highly proliferative subset of interstitial cells (i-cells),

which are bipotent neuronal progenitors continuously providing apical neurons and nematoblasts (sensory mechanoreceptor precursors; Miljkovic-Licina *et al.*, 2007). In an evolutionary perspective it is worth highlighting that also in mice Gsx homologs Gsh1 and Gsh2 appear to control the size and the identity of neuronal progenitor pools (Toresson and Campbell, 2001; Yun *et al.*, 2003), as later described. Analogously in *Nematostella* anthox2 seems to be involved in apical neurogenesis (Galliot and Quiquand, 2011).

1.9.2 Gsx in protostomes: Lophotrochozoa

Lophotrocozoans (mainly represented by platyhelminthes, annelids and mollusks) are one of the major *superphylum* constituting the protostomes clade, together with *Ecdysozoa*.

Studies on *Gsx* (and in general on ParaHox genes) in this group are mostly related to annelids (*phylum Annelida*) and solely based on expression profile analyses, in respect to which the function of the gene is hypothesized. One of the first works in these organisms is about *Capitella teleta*, in which *CapI-Gsx* transcript is really circumscribed, both in space and time. It is indeed only transiently expressed during early stages of brain formation, in a subset of anterior neuroectoderm cells (Fröbius and Seaver, 2006).

Much more interesting, instead, are data coming from the probably best studied worm in this context, *Platynereis dumerilii*, where *Pdu-Gsx* expression is very dynamic. Since the pre-larval stage (24hpf) the gene is present in the prospective neural tissue and later during the differentiation of the trunk CNS. In particular *Pdu-Gsx* is present in a subset of neurons which corresponds to the intermediate zone of vertebrate and Drosophila neural plate (Denes *et al.*, 2007), a correspondence that is interesting from the evolutionary point of view, as later described. Furthermore, *Pdu-Gsx* is not only restricted to neural regions, being also expressed at larval stage in two bilateral cell clusters in the stomodeum (mouth precursor) and later on in small cell clusters in the midgut and posterior foregut (Hui *et al.*, 2009).

Dual domains of expression, in the CNS and the gut, are similarly observed in another polychaetes, *Nereis virens* (Kulakova *et al.*, 2008), and also outside the annelids, in *Gibbula varia* (*phylum Mollusca*; Samadi and Steiner, 2010).

1.9.3 Gsx in protostomes: Ecdysozoa

Ecdysozoa is the second *superphylum* of protostome animals, constituted by *Artropoda*, *Nematoda* and other smaller *phyla*. In this case most of the data about ParaHox genes pertain the fruit fly *Drosophila melanogaster*. Differently from what has been illustrated up to now, a precise and certain role has been attributed to *Gsx* ortholog in insects, called *ind*, during *Drosophila* embryonic development, not only on the basis of its expression but also through deep molecular studies that allowed to reconstruct the regulatory network in which the transcription factor is involved. In 1998 *ind* was isolated and thus recognized as one of the key players in dorsoventral (DV) patterning of *Drosophila* nervous system (Weiss *et al.*, 1998).

During the development of CNS in *Drosophila* the neuroblasts (neural stem cells) form an orthogonal grid of four rows along the anterior-posterior axis of the embryo and three columns (ventral, intermediate and dorsal) along the dorsoventral one, so that each one of them acquires a unique fate based on its position and time of formation. As one can easily suppose, a complex genetic hierarchy is involved in this specification and many works have been devoted to its definition. In the trunk neuroectoderm, the first genes identified in the establishment of the dorsoventral identity of the three columns were *vnd* (ventral nervous system defective), *ind* (intermediate neuroblasts defective) and *msh* (muscle segment homeobox), respectively expressed in ventral, intermediate and dorsal column. The three genes are involved in a repressive cascade whereby Vnd represses *ind* transcription in the ventral column, while Ind in turn represses *msh* in intermediate column, in both cases through a direct interaction with the enhancer of the regulated gene (***figure 1.10***. Mc Donald *et al.*, 1998; Weiss *et al.*, 1998; Von Ohlen *et al.*, 2009). In summary through this genetic hierarchy the more

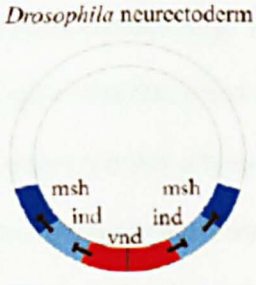


Figure 1.10 Interactions between ind, msh and vnd factors.
Schematic cross-section diagram showing expression of *msh*, *ind* and *vnd* in the *Drosophila* neuroectoderm. Uni-directional repressive interactions are indicated by black lines. (Adapted from Winterbottom *et al.*, 2010).

ventrally expressed gene establishes, *via* transcriptional repression, the ventral limit of the gene expressed immediately dorsal to it.

These initial findings have been of course more and more deepened and in particular, for what concerns Ind, many data are now available about its function, structural organization and transcriptional regulation.

First of all, it has been demonstrated that Ind exerts its function as transcriptional repressor, although some data point to the presence of an activation domain in its sequence (Von *Ohlen et al*, 2007a; Von Ohlen *et al.*, 2007b).

Ind repressor activity is mainly exerted through the interaction with known co-factors, such as Groucho (Gro). As a co-repressor, Gro does not bind DNA directly, but it is anchored to DNA only *via* conserved polypeptide sequences present in the transcriptional regulators it interacts with. One of these sequences is called Engrailed homology 1 (Eh1), predominantly present in the homeodomain-containing TFs, just such Ind (Smith and Jaynes, 1996). Although the Eh1 domain is quite variable, a few conserved amino acids can be recognized: Phe-x-Ile-x-x-Ile, (F-x-I-x-x-I) where x is any amino acid (Goldstein *et al.*, 2005).

Comparative sequence analyses combined with functional studies led to the identification in Ind protein sequence of at least two repressive domains, both at the N-terminal. One of them is the conserved Eh1, while the second one, called PstI, seems to be insects specific, since it is not conserved in vertebrates at the sequence level; moreover PstI bears no significant homology to any known functional domain (Von Ohlen *et al.*, 2007a; Von Ohlen and Moses, 2009). It has been clearly demonstrated that Eh1 acts also in Ind as a repressive domain through the binding with Groucho, whereas PstI exerts a repressive function in a Groucho independent manner. Both domains are required to confer the maximal repressional activity to the protein (Von Ohlen and Moses, 2009).

Besides Eh1 and PstI, the authors also identified a third domain, which is an activator domain. This is not much surprising since, in addition to its function as a transcriptional repressor, Ind is

also able to in maintain its own expression during *Drosophila* embryonic development, acting in this case as a transcriptional activator (Von Ohlen *et al.*, 2007b).

As previously mentioned, one of the target of Ind transcriptional repression in *Drosophila* is the homeobox gene *msh*. More recently other Ind regulated genes have been identified, amongst which of particular interest is the *Pax6 Drosophila* ortholog, *ey*. A cross-talk between Ind and ey has been, indeed, demonstrated in a way that the two genes are able to repress each other (Von Ohlen *et al.*, 2007a).

The regulatory pathways so far described are mostly referred to the truncal neuroectoderm, that leads to the formation of the ventral nerve cord. A bit less is known about the mechanisms controlling pattern formation of *Drosophila* brain. Interestingly, in this region of the embryo *vnd*, *ind* and *msh* exhibit specific and different genetic interaction. It has been indeed recently demonstrated that the regulatory network underlying the formation of the *Drosophila* brain requires mutual repressive mechanisms, rather than a one-way cascade. In particular, the expression pattern of *vnd*, *ind* and *msh* is much more complex and dynamic during development. In this genetic circuit, for example, *msh* is repressed by Ind, but it is in turn able to block *ind* transcription; the same is true for vnd/msh and vnd/ind interactions. Therefore, the dorsoventral patterning of the procephalic neuroectoderm is a consequence of the net of activation/repression mechanisms between the different involved factors, in different regions of the embryo and in a certain embryonic stage (Seibert and Urbach, 2010).

In an evolutionary perspective, the identification of these bidirectional cascades is very interesting, since it somehow recalls the repressive relationship between pairs of homeodomain proteins that are responsible for the DV patterning of the vertebrate neural tube (Seibert and Urbach, 2010).

For what concerns the mechanisms controlling *ind* transcriptional regulation, the data point to Dorsal and EGFR (Skeath, 1998; Weiss *et al.*, 1998; Von Ohlen and Doe, 2000; Hong *et al.*, 2008) as key regulators of *ind* expression in the intermediate column. Dorsal is the more potent activator of *ind*, primarily acting through a direct binding on *ind* enhancer. Besides that, an

indirect mechanism occurs: in the intermediate column Dorsal induces the EGFR signaling, which is in turn able to activate *ind*, relieving its repression due to the transcriptional repressor Cic (Capicua; Jiménez *et al.*, 2000; Von Ohlen and Doe, 2000; Lim *et al.*, 2013).

1.9.4 Gsx in non-chordate deuterostomes: Ambulacraria

Ambulacraria is a clade of deuterostomes non-chordate animals, constituted by two major *phyla*, *Echinodermata* (e.g. starfish, sea urchins, sea cucumbers) and *Hemichordata*.

Studies regarding ParaHox genes in the echinoderms are focused on their expression pattern in the sea urchin *Strongylocentrotus purpuratus* (Arnone *et al.*, 2006) and the sea star *Patiria miniata* (Annunziata *et al.*, 2013).

Unfortunately, little can be deduced about *Gsx* in *Patiria miniata*, since *in situ* hybridization experiments revealed that its expression is really poor during development, hardly above the minimum level detectable (Annunziata *et al.*, 2013). Conversely, for the echinoid *S. purpuratus* a role of Gsx in the developing nervous system can be presumed, given its presence from gastrula through pluteus stages in a small ectodermal domain, probably neural (Arnone et al., 2006).

Among the hemichordates, the ParaHox cluster has been very recently identified in *Ptychodera flava* (Ikuta *et al.*, 2013). Here too only expression data are available which indicate that as for echinoderms, *Gsx* is restricted to a small territory in the ectoderm, in particular in few cells around the blastopore at gastrula stage, disappearing at larval stage.

1.9.5 Gsx in chordate deuterostomes: vertebrates

In the "jump" from protostomes to deuterostomes we witnessed to a simplification or, more precisely, to a reduction of the *Gsx* expressing territories in the different analyzed species. This is absolutely false in the case of vertebrates, where the scenario is probably even more complicated by the presence of two *Gsx* paralogs, named *Gsh1* and *Gsh2* (or *Gsx1* and *Gsx2* too).

Gsh1, *Pdx* (Xlox ortholog) and *Cdx2* (one of the three Cdx paralogs in vertebrates) constitutes the intact vertebrate ParaHox cluster previously mentioned, whereas the remaining genes exist in degenerate clusters containing single ParaHox genes (Ferrier *et al.*, 2005; Illes *et al.*, 2009). Both expression and functional studies, mainly conducted on *Xenopus tropicalis* and mouse, have been performed, demonstrating that also in this case the principal role of Gsh1/2 is related to CNS development.

In *Xenopus tropicalis Gsh1* and *Gsh2* exhibit an highly dynamic expression pattern within the developing nervous system since early neural plate stage and throughout development, in the forebrain, midbrain, hindbrain and spinal cord, with a nice bilaterally symmetrical distribution across the mid-line. In particular, among the territories in which *Gsh1* transcript can be detected, is worth mentioning pretectum and tectum (two structures involved in processing visual and auditory stimuli), thalamus and hypothalamus, olfactory bulb and a cell population giving rise to interneurons. *Gsh2* is more or less expressed in the same territories, even if regions of perfect co-localization with *Gsh1* are quite rare. Of particular relevance is the presence of the gene also in the endoderm, as occurs in more basal organisms, probably providing a further evidence about the ancestral role of Gsx (see later. Illes *et al.*, 2009).
In mouse, the two genes specifically mark the CNS, with a bilateral distribution similar to the one described for *Xenopus* and *Drosophila*. Interestingly *Gsh1* is also expressed in the optic stalk (optic nerve primordium. Hsieh-Li *et al.*, 1995; Valerius *et al.*, 1995).

A few expression data are also available for *Gsh1* in medaka fish and zebrafish. In both cases the gene is expressed with a dynamic pattern during development, in many regions of the central nervous system, among which is worth mentioning again hypothalamus primordium, spinal cord (in a region that in zebrafish will generate interneurons) and optic tectum of medaka fish (Deschet *et al.*, 1998; Cheesman and Eisen, 2004).

Given this complex picture, it is easy to imagine that *Gsh1* and *Gsh2* play multiple roles in CNS development, probably each one of them through a distinct non-redundant function. Firstly, all the data collected so far indicate that in vertebrates Gsh1 and Gsh2 act as transcriptional repressors, as in *Drosophila*, and probably with the help of Groucho, while, differently than in *Drosphila*, no evidence are available about a putative activator domain. In general they have been associated to three main processes, not necessarily independent. They promote the dorsoventral patterning along the neural tube (Toresson *et al.*, 2000; Winterbottom *et al.*, 2010) and are also able to specify neuronal progenitors (Toresson and Campbell, 2001; Yun et al., 2003; Winterbottom *et al.*, 2011); furthermore they regulate the balance between proliferation and differentiation of neuronal precursor cells during development (Toresson and Campbell, 2001; Pei *et al.*, 2011).

Winterbottom and co-workers, for example, defined the regulatory mechanism in which Gsh2 is involved for the specification of interneurons during primary neurogenesis in *Xenopus*. Even if the relative domains of expression of *Gsh2*, *Nk6* (related to *vnd* in *Drosophila*) and *Msx* (ortholog of *msh* in *Drosophila*) are comparable between *Xenopus* and *Drosophila* neural plates (***figure 1.11 A***, compated to ***1.10***), their interactions are totally different, starting from the observation that *Msx1* is not regulated by Gsh2. Furthermore, in the frog, Gsh2 is able to promote differentiation of primary interneurons through the direct repression of two different homeobox factors, *Dbx* and *Iro3* (Winterbottom *et al.*, 2010; Winterbottom *et al.*, 2011).

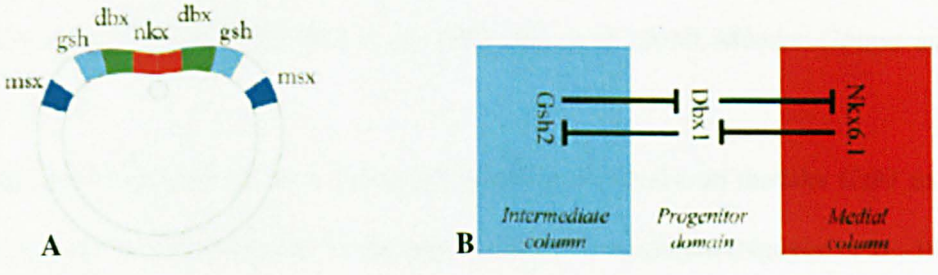


Figure 1.11 Interactions between Gsh2, Msx and Nk6.1 factors.
A. Schematic cross-section diagrams showing relative positions of *Nkx*, *Gsh*, *Dbx* and *Msx* in the *Xenopus* open neural plate. **B.** Proposed model of bidirectional repressive interactions between Gsh2, Dbx1 and Nkx6.1 transcription factors in the *Xenopus* open neural plate. (Adapted from Winterbottom et al., 2010)

In addition, the authors provide evidences of a bidirectional repressive network including Nkx6/Dbx/Gsh2 that patterns the medial region of the neural plate (***figure 1.11 B***), in a way that is more similar to the brain patterning in *Drosophila*, rather than the specification of the trunk intermediate column (Winterbottom *et al.* 2010).

Among vertebrates, many examples of the role of Gsx in DV patterning of the nervous system are available also in mouse. In this context, one of the most studied case is the opposite role exerted by Pax6 and Gsh2 in the control of the boundary between the cortex and the Lateral Ganglionic Eminence (LGE) in the telencephalon (Toresson *et al.*, 2000; Yun *et al.*, 2003; Carney *et al.*, 2009). In the embryonic telencephalon indeed, the cortex (cerebral cortex precursor) and the LGE (striatum precursor) are two adjacent territories, specifically expressing a number of developmental control genes. In particular Gsh2 in required to repress *Pax6* in the LGE, conversely Pax6 is necessary (even if not sufficient) for the repression of *Gsh2* in the cortical progenitors (Toresson *et al.*, 2000). Interestingly this data remind the crosstalk between ey and Ind in *Drosophila* (Von Ohlen et al., 2007a), thus suggesting a kind of conservation in the function of these genes in animal phylogeny.

It has also been demonstrated that both *Gsh1* and *Gsh2* genes are not only required for the patterning of LGE progenitors, but also for the control of their proliferative characteristic, acting in the opposite way: Gsh2 maintains telencephalon progenitors in an undifferentiated state, whereas Gsh1 promotes progenitor maturation toward neurogenesis (Toresson *et al.*, 2000; Toresson and Campbell, 2001; Yun *et al.*, 2001; Pei *et al.*, 2010; Méndez-Gómez and Vicario-Abejón, 2012).

The latter could intriguingly be a conserved property, since also in medaka *Gsh1* expression is associated with proliferating cells in the optic tectum and hindbrain (Nguyen *et al.*, 1999).

1.9.6 Gsx in chordate deuterostomes: cephalochordates and urochordates

The first evidence about the ParaHox cluster, as already mentioned, came from the amphioxus *B. floridae*. *AmphiGsx* is expressed in two temporally distinct domains (early and late domain),

in both cases included in neural regions. The early domain, during neurulation, consists of four cells in the neural tube (partially overlapping with *AmphiXlox* expression), whereas the late domain is located more anteriorly, in the cerebral vesicle (anatomical homologue of the vertebrate forebrain/midbrain), in both left and right sides along the AP axis (Brooke *et al.*, 1998; Osborne *et al.*, 2009). Concerning its transcriptional regulation, no data have been reported so far beside the evidence that Amphi-ParaHox cluster is regulated by the retinoic acid (RA), a signalling cascade already demonstrated to be involved in the establishment of chordate AP axis, mainly *via* Hox genes. *AmphiGsx*, in particular, expands its domain anteriorly after RA treatments, yet confined in the neuroectoderm (Osborne *et al.*, 2009). This regulation could occur directly through RAREs sequences (RA response elements) that have been identified in the cluster.

Concerning Urochordates the expression pattern of *Gsx* has been described in the ascidian *C. intestinalis* (Hudson and Lemaire, 2001; ***figure 1.12***).

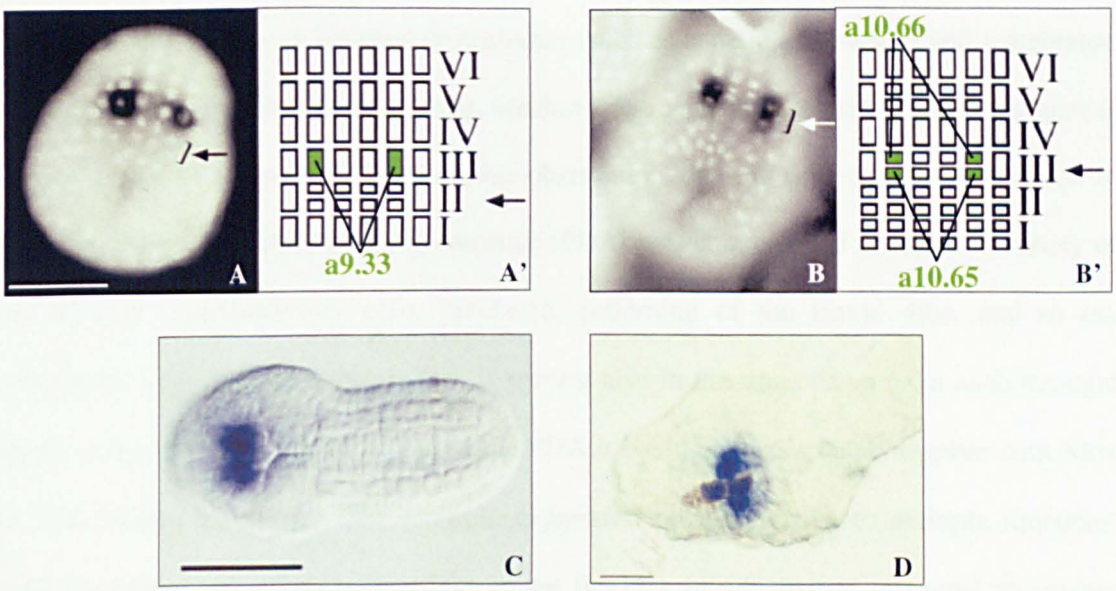


Figure 1.12 Expression pattern of *Ci-gsx*. Embryos at different developmental stages stained for *Ci-gsx* probe. In each case dorsal views are shown and the scale bar is 50 µm. **A, B:** *In situ* hybridization of gastrula-neurula stage embryos co-stained with a nuclear dye to identify cell positions. Arrows mark rows that have undergone mitosis. **A', B':** Schematic representation of the neural plate. Roman numeral indicate the different rows. Arrows indicate newly cleaved cells. Cells that express *Ci-gsx* are highlighted in green and their names are indicated. **C.** Early tailbud stage, anterior is to the left. **D.** Late tailbud embryo, stained for tyrosinase (brown) too. Anterior is to the left. (Hudson and Lemaire, 2001)

Ci-gsx appears quite early during *Ciona* development, being clearly detectable since neural plate in the a9.33 blastomere pair and later in the descendant a10.65 and a10.66 cell couples (***figure 1.12 A-B'***). From early tailbud stage, expression is located in two bilateral domains (***figure 1.12 C***) that in late tailbud embryos are included in two regions of the developing brain vesicle, one of which lies around the ocellus (***figure 1.12 D***).

As described above, in *Ciona* the right a9.33 and a9.37 pairs of the neural plate stage, has been proposed to originate the photoreceptor cells of the ocellus (Cole and Meinertzhagen, 2004). Therefore, the presence of *Ci-gsx* in these cells, at the stage when their fate start to be restricted, could possibly suggest the idea of its putative involvement in the early specification of this lineage.

1.10 To sum up...

The data described so far show that Gsx is widely distributed along animal phylogeny, probably exerting its function mainly in the nervous system, through the specification of several cell types and/or the control of their differentiation state. Similarities can be recognized in the width of the expression territories between protostomes (such as annelid and insects) and vertebrates, passing through some cases (echinoderms, urochordates, cephalochordates) in which the gene is restricted in a smaller domain. Based on this observation, one can suppose that the function of Gsx in the Protostome-Deuterostome Ancestor (PDA) was complex and related to a variety of roles in eyes, neurosecretory cells, hindbrain, patterning of the neural tube, and so on. Furthermore, since in a few species Gsx is present also in the stomodeum or in endoderm/gut regions, it is possible to speculate that in the PDA it could also contribute, together with Xlox and Cdx, to gut patterning. This multiple expression pattern, related to multiple functions, would have been secondarily simplified in the lineages in which it is restricted to smaller patches (Hui *et al.*, 2009).

In order to verify this hypothesis, molecular data from as many species as possible are necessary, to outline a clear picture of the functions of this ancient gene during evolution. Our

model system *Ciona intestinalis* exhibits all the required properties to represent a useful joining link for this story.

AIM OF THE THESIS

The main aim of my PhD thesis has been to explore the putative role of Gsx in the early specification of photoreceptor cells in *Ciona intestinalis*, adding new information about the functions of this transcription factor in the developing nervous system in a model system closely related to vertebrates. Besides few expression patterns (for example in amphibians, mouse, fishes), no functional data, indicating an involvement of Gsx in photoreceptor cells formation, have been reported so far.

Ci-gsx is known to be expressed since gastrula stage in the right a9.33 and a9.37 cells of the neural plate, proposed to be photoreceptor cells precursors. Based on these preliminary observations, the first goal of my PhD work has been the refinement of *Ci-gsx* expression pattern during *Ciona intestinalis* development through detailed double *in situ* hybridization experiments, in order to confirm its expression in photoreceptor cells precursors.

Afterwards, the isolation of a genomic fragment (*pGsx*) containing *Ci-gsx* regulatory region has been instrumental for the two next objectives of my project: exploring Ci-gsx function in photoreceptor cells formation and studying the molecular mechanisms required for its transcriptional regulation.

The experimental strategy adopted consisted in promoter-guided perturbation strategies, thanks to which chimeric forms of Ci-gsx, that should be able to interfere with the activity of the endogenous gene, were specifically directed in *Ci-gsx* territories by using its own promoter (*pGsx*). At the same time, I dissected the *pGsx* regulatory region in order to identify the minimal *Ci-gsx* promoter and the putative *trans*-acting factor(s) responsible for *Ci-gsx* regulation (activation/repression).

Chapter 2

MATERIALS AND METHODS

2.1 *Ciona intestinalis* eggs and embryos collection

Ciona intestinalis adults were collected in the bay of Naples or of Taranto, and maintained at the Stazione Zoologica A. Dohrn (Naples) Aquaculture Service under constant illumination, in sea water tanks equipped with appropriate water circulation and filtering system. Ripe oocytes and sperm were collected surgically and kept separately until in vitro fertilization. Fertilized eggs were used in transgenesis and/or *in situ* hybridization experiments. Embryos were raised in Millipore-filtered sea water (MFSW) at 18-20°C.

2.2 Chemical dechorionation and *in vitro* fertilization

Before to perform electroporation of the fertilized eggs, it is necessary to deprive the eggs of the chorion. The chemical dechorionation has been effectuated in a Petri dish with a thin layer of 1% agarose in MFSW, putting the eggs for 5-6 minutes in a pH 10 solution of Thioglycolic acid (1%) and Proteinase E (0.05%) in MFSW. During this time, the eggs have been shaken continuously in this solution, using a glass pipette, to remove the chorion and the follicular cells surrounding the eggs. After this step the eggs have been washed several times to remove the residual solution and then fertilized *in vitro* with sperm collected from two or more individuals to avoid self-sterility problems. After 10 minutes, fertilized eggs have been washed 2-4 times to eliminate the exceeding sperm; at this point they have been used for transgenesis experiments or, alternatively, the embryos have been grown in MFSW at 18-20°C, and fixed at the suitable stages to perform whole-mount *in situ* hybridization (WMISH).

2.3 Transgenesis *via* electroporation

The fertilized eggs have been immediately transferred in a solution containing 0.77 M Mannitol and 50-100 μg of DNA. The electroporation has been carried out in Bio-Rad Gene Pulser 0.4 cm cuvettes, using a Bio-Rad Gene Pulser II electroporator , at constant 50 V and 500-800 μF, in order to have a time constant of 14-20 m/seconds. The embryos have been allowed to develop

until the desired stage on 1% agarose coated Petri dishes, at 18-20°C. Depending on the electroporated constructs and on the purposes of the experiment, the embryos have been fixed for Whole Mount *in situ* Hybridization or analyzed at the microscope.

In order to be sure of the electroporation success, only the experiments in which at least 60% of the embryos developed normally were selected for my analyses. All the constructs have been tested in at least three different batch of animals; percentages reported in the results have been calculated taking in to consideration at least 100 embryos for each construct.

2.4 Embryos observation and imaging analyses

For the observation at the microscope of fluorescence and phenotypes, embryos and larvae have been observed *in vivo*. To avoid embryos movement, late tailbud and larvae have been sedated using menthol crystals in the sea water. The embryos have been then placed on a microscope slide; a cover slide with some plasticine at the corners have been positioned on the top of a sea water drop, containing the embryos, and pressed until the volume resulted uniformly distributed. DIC and fluorescent images have been taken with a Zeiss Axio Imager M1 microscope equipped with an Axiocam digital camera.

For confocal images, embryos have been analyzed with a Zeiss confocal laser scanning microscope LSM 510.

2.5 PCR amplification from genomic or plasmid DNA

The PCR reactions have been performed in a total volume of 50 µl, using about 100 ng of DNA as template, 0.2 mM of dNTP mix (dATP, dTTP, dCTP, dGTP), 1x PCR buffer (Roche), 0.05U/µl of Taq DNA polymerase (Roche) and 2 pmol/µl of each forward and reverse suitable oligonucleotides. The PCR amplification program has been set as follows.

- First step (1 cycle). DNA denaturation: 5' at 95°C.
- Second step (repeated for 35 cycles).

  DNA denaturation: 1' at 95°C.

Oligonucleotides annealing: 2' at the suitable temperature for plasmid DNA, 4' at suitable temperature for genomic DNA (the temperature used in this step has been set at least 5-8°C below the melting temperature of the oligonucleotides).

Polymerization: 72°C for a suitable time, calculated considering the desired amplified fragment length and the Taq DNA Polymerase processivity, that is around 1 Kb/minute.

- Final elongation step: 10' at 72°C.

The amplified fragments have been separated from the template DNA and from dNTPs and oligonucleotide excess by gel electrophoresis using, as fragment length marker, 1x GeneRuler™ 1Kb DNA Ladder (Fermentas), 1x GeneRuler™ 100 bp DNA Ladder (Fermentas) or 1x Lambda DNA/HindIII, 2 (Fermentas), according to the expected length of the fragment.
The fragments have been isolated and purified by gel extraction (GenElute™Gel Extraction Kit, Sigma). The concentration have been evaluated by gel electrophoresis using as marker 1x Lambda DNA/HindIII, 2 (Fermentas).

2.6 DNA gel electrophoresis

Preparative and analytic DNA gel electrophoreses have been performed on 0.8%, 1% or 1.5 % of agarose gel in 1x TAE buffer (Stock solution 50x: 252 g of Tris base; 57.1 ml glacial cetic acid; 100 ml 0.5 M EDTA; H_20 to 1 liter), considering the length of the DNA to be run and adding 0.5 µg/ml Ethidium Bromide (EtBr).

2.7 DNA gel extraction

Digested and PCR amplified fragments have been extracted from gel cutting them with a sterile sharpen blade, using the GenElute™ Gel Extraction Kit (Sigma-Aldrich), following the manufacturer's instructions. After the extraction, the concentration has been estimated by gel electrophoresis.

2.8 DNA digestions with restriction endonucleases

Analytic and preparative plasmid DNA digestions have been performed with the appropriate restriction endonucleases in total volumes of at least 20 times more than the enzyme volume used. The digestion reaction has been prepared as follows: the solution contained the desired amount of DNA, suitable restriction enzyme buffer (1/10. Roche; New England Biolabs; Amersham), restriction enzyme/s (5 units enzyme per 1 μg of DNA) and BSA (1/100, if required). Reaction specific temperatures have been used as suggested by manufacturer's instructions.

2.9 Fill in of 5' overhangs

A "fill-in" reaction is used when is necessary to create blunt ends on fragments created by cleavage with restriction enzymes that leave 5'overhangs. I applied this strategy to ligate restriction products whose termini are not compatible, when I had no chance to use endonucleases that produce compatible overhangs. Fill-in reaction requires the Klenow Fragment, a proteolytic product of *E. coli* DNA Polymerase I, which retains polymerization and 3'-5' exonuclease activity, but has lost 5'-3' exonuclease activity of the native enzyme (Jacobsen *et al.*, 1974). The enzyme catalyzes the addition of mono-nucleotides from deoxynucleoside-5 '-triphosphates to the 3'-hydroxyl terminus of a primer/template DNA. This property is used to synthesize DNA complementary to single-stranded DNA templates, since Klenow Fragment retains the polymerization fidelity of the holoenzyme.

To set up this reaction a convenient amount of digested DNA is incubated in the following reaction mix: 4 mM dNTP mix (dATP, dTTP, dCTP, dGTP), filling buffer 1x, 1 U of Klenow enzyme (Roche) for each μg of DNA, in a total volume of 20 μl. The reaction is then incubate at 37°C for 15', and inactivated by heat at 65°C for 10'.

2.10 DNA dephosphorylation

In order to prevent self-ligation, a convenient amount of double strand linearized DNA has been incubated with 1U of Calf Intestinalis Alkaline Phosphatase enzyme (CIAP; Roche) per 1 pmol

5' ends of linearized DNA, in 1x CIP dephosphorylation buffer (Roche), at 37°C for 30'. After this time, a second aliquot of CIAP has been added, and the reaction has been carried on for another 30', at 37°C. Subsequently, the dephosphorylated DNA has been purified by gel extraction (see paragraph 2.7).

2.11 DNA ligation

Each ligation reaction has been carried out in a final volume of 20 µl mixture containing 1x T4 Ligase buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, pH 7.5) and 1 µl of T4 DNA Ligase (New England Biolabs) at 1U/ µl. The proportion of plasmid vector and insert DNA was usually kept 1:4, and the total amount of DNA was kept within 50-100 ng. The reaction mix has been incubated at 16 °C overnight or 1,5 hour at R.T., and used to transform competent bacteria.

2.12 Bacterial cells electroporation

By this approach it is possible to transform bacterial cells with plasmids containing DNA of interest. Briefly, the circular plasmid DNA and competent *E. coli* bacterial cells (prepared by the Molecular Biology Service of the Stazione Zoologica A. Dohrn in Naples), were placed in a 0.2 cm electrocuvette. The electrocuvette was subjected to an electric pulse at constant 1.7 V using a Bio-Rad Gene Pulser™ electroporation apparatus.

The transformed *E. coli* cells were allowed to recover for one hour at 37°C in 1ml LB medium (NaCl 10g/l, bactotryptone 10g/l, yeast extract 5g/l,). An aliquot was spread on a pre-warmed LB solid medium (NaCl 10g/l, bactotryptone 10g/l, yeast extract 5g/l, agar 15g/l) in the presence of specific selective antibiotic grown at the same temperature overnight.

2.13 PCR screening

It is possible to carry out a PCR reaction using as template a single bacterial colony and in the same time grow the colony. Half of each single colony was placed in a PCR mixture described

below, and half grown in 3 ml of LB liquid medium in the presence of the suitable antibiotic (50μg/ml), 8-12 hours shaking at 270 rpm, at 37 °C.

The PCR reactions have been set in a total volume of 20 μl, with the following composition: 1x PCR buffer (Roche); 0.2 mM dNTP mix (dATP, dTTP, dCTP, dGTP); 1 pmol/μl of each forward and reverse suitable oligonucleotides; and 0.025 U/μl Taq DNA polymerase (Roche; Biogem). PCRs have been carried out with the following protocol:

- First step (1 cycle). DNA denaturation: 5' at 95°C.
- Second step (repeated for 30 cycles).

  DNA denaturation: 1' at 94°C.

  Oligonucleotides annealing: 1' at the suitable temperature (according to the melting temperature of oligonucleotides)

  Polymerization: 72°C for a suitable time (1 min/kb).

By gel electrophoresis analysis, the samples presenting a band of expected sixe have been identified and plasmid DNA has been purified from the corresponding bacterial colonies.

2.14 Plasmid DNA Mini- and Maxi-preparation

A single bacterial colony containing the plasmid DNA of interest was grown in a suitable volume of LB (4-5 ml for Mini-preparation, 200-400 ml for Maxi-preparation) in the presence of the appropriate antibiotic shaking at 37°C overnight. The Sigma-Aldrich Plasmid Purification kit, based on alkaline lyses method, was used to isolate the plasmid DNA from the cells according to the manufacture's instruction.

2.15 Sequencing

The DNA sequences have been obtained using the Automated Capillary Electrophoresis Sequencer 3730 DNA Analyzer (Applied Biosystems, Foster City, CA) by the Molecular Biology Service of the Stazione Zoologica A. Dohrn in Naples.

2.16 Oligonucleotides synthesis

All used synthetic oligonucleotides were prepared with a Beckman SM-DNA Synthesizer at the Molecular Biology Service of the Stazione Zoologica A. Dohrn in Naples.

2.17 Digested plasmids purification

To eliminate protein contaminations, the plasmid DNA linearized in order to obtain the template for riboprobes synthesis has been purified with 1 volume of phenol:chloroform:isoamylic alcohol (25:24:1), vortexed vigorously and centrifuged at 13000 rpm for 5 minutes at 4°C. The soluble phase has been recovered and 1 volume of chloroform:isoamylic alcohol (24:1) has been added; the sample has been vortexed vigorously and centrifuged at 13000 rpm for 5', at 4°C. The aqueous phase has been recovered and the DNA has been precipitated adding 3 volumes of ethanol 100% and 1/10 volumes of Sodium acetate 3M pH 5.2. The sample has been mixed and stored over night at -20°C or 1 hour at -80°C. Then, it has been centrifuged at 13000 rpm for 15', at 4°C. The precipitated DNA has been washed with ethanol 70% (sterile or DEPC-treated), centrifuging at 13000 rpm for 15' at 4°C. The ethanol has been removed and the sample has been air-dried at R.T. At the end, the DNA has been diluted in a suitable volume of H2O (sterile or DEPC-treated), and its concentration has been evaluated by gel electrophoresis, using 1X Lambda DNA/Hind III Marker 2 (Fermentas), and using a spectrophotometer (Nanodrop 1000, Thermo SCIENTIFIC), reading the values at the wavelengths of 230, 260 and 280 nm and calculating the ratio between 260/230 nm and 260/280 nm to ascertain respectively the absence of chemical (phenol, ethanol) and protein contamination.

2.18 Ribonucleic probes preparation

2.18.1 RNA labelling

The plasmid, containing the template to be transcribed, has been conveniently digested and purified, as described in paragraph 2.17. 1 µg of purified, linearized DNA has been used as template for the ribonucleic probe synthesis. This template has been added to the following reaction mix: transcription buffer (1/10; Roche); Digoxigenin or Fluorescein labelling mix,

containing 1 mM of ATP, CTP and GTP, 0.65 mM UTP and 0.35 mM DIG-11-UTP or 0.35 mM fluorescein-12-UTP (Roche); Sp6 or T7 RNA polymerase (2U/µl; Roche); Protector RNase inhibitor (1U/ µl).

The reaction has been performed in a total volume of 20 µl (in H_2O DEPC-treated). The mix has been briefly centrifuged and incubated for 2 hours at 37°C. Then, DNaseI RNase free (0.9U/µl) has been added in order to remove the DNA template. The sample has been incubated for 20' at 37°C. Finally, the reaction has been stopped adding EDTA pH 8.0 (16 mM). The synthesized ribonucleic probes have been purified using the mini RNeasy mini kit (QIAGEN), following manufacturer instructions. The ribonucleic probe quality has been checked by gel electrophoresis and the concentration quantification has been evaluated by Dot Blot analysis (see the next paragraph). One volume of deionized formamide has been added to the recovered samples, immediately stored at -80°C till the use.

The ribonucleic probes listed (**table 2.1**) have been synthesized starting from cDNA clones present in N. Satoh *C. intestinalis* gene collection 1, available in the laboratory.

Table 2.1 Genes of which ribonucleic probes have been synthesized

Gene name	Corresponding clone in N. Satoh gene collection 1	Probe lenght
Ci-gsx	citb029c24 (plate: R1CiGC31m18)	1250 bp
Ci-Six3/6	cicl021e08 (plate: R1CiGC11m13)	1500 bp
Ci-Meis	citb035l22 (plate: R1CiGC32a16)	1995 bp
Ci-Ptf1a	cilv050i16 (plate: R1CiGC44e22)	1000 bp
Ci-msxb	cign067l18 (plate: R1CiGC42h24)	1028 bp
Ci-SoxB1	ciad002g15 (plate: R1CiGC01k09)	2550 bp
Ci-Rx	Dr Branno's gift, (D'Aniello et al., 2006)	2500 bp

2.18.2 Ribonucleic probes quantification by Dot Blot analysis

The concentration evaluation of the DIG-11-UTP or fluorescein-12-UTP incorporated in the ribonucleic probes has been estimated making serial dilutions of the sample ribonucleic probes and of a Control RNA of reference (Roche), in a buffer containing DEPC-treated H_2O, 20x SSC, formaldehyde (5:3:2). 1 µl of each dilution has been blotted on a membrane Hybond-N (Amersham) and air-dried at R.T. The RNA has been UV-crosslinked on the membrane with a Stratalinker for 30". The membrane, with the UV-cross-linked RNA on it, has been washed in blocking solution (5% BSA in 0.1 M maleic acid pH 7.5), for 30', shaking at R.T. Subsequently,

the membrane has been incubated with anti-DIG or anti-Fluo phosphate alkaline antibody (0.15 U/ml; Roche) in blocking solution for 1 hour, shaking at R.T. To remove unbound antibodies, the membrane has been washed twice in a solution containing 0.1 M maleic acid pH 7.5 and 0.15 M NaCl for 15', at R.T. The membrane has been equilibrated in the detection solution (100 mM NaCl; 100 mM Tris pH 9.5; 50 mM MgCl2, in H2O) for 5', at R.T., and then incubated in the dark in the same solution in which Nitro-Blue Tetrazolium Chloride (NBT; 35 µg/ml) and 5-Bromo-4-Chloro-3'-Indolylphosphate p-Toluidine (BCIP; 175 µg/ml) have been added. The reaction has been monitored every 4-5' and blocked at the suitable moment, putting the membrane under running water. The membrane has been dried on 3MM paper and the concentration of the DIG-11-UTP or fluorescein-12-UTP incorporated in the ribonucleic probes has been calculated from the comparison with the control RNA spots.

2.19 Whole Mount In situ Hybridization (WMISH) assays

2.19.1 Embryos preparation

Wild type or transgenic embryos at suitable embryonic stages have been fixed at R.T. 90' or at 4°C over night, in a mixture containing 4% paraformaldehyde, 0.1 M MOPS pH7.5, 0.5 NaCl. Subsequently they have been washed three times in PBS 1x and dehydrated in graduated scale of ethanol (30% - 50% - 70% ethanol in distilled water DEPC-treated). They have been stored at -20°C until used.

2.19.2 WMISH protocol for single and double in situ

Day 1. The dehydrated stored embryos have been firstly re-hydrated in a graduate scale of 100% - 70% - 50% - 30% methanol in PBT (that is PBS + 0.1% tween 20): one wash of 20' for each methanol solution.

After that, the samples have been washed 3x15' in 1 ml PBT at R.T. and incubated 30' at 37°C in 1 ml PBT containing 2 µg/ml protease K for dechorionated embryos or 4 µg/ml for non dechorionated embryos, to increase the permeability to the cells and accessibility to mRNA target. The reaction has been stopped by a wash in 2mg/ml glycine in PBT. After digestion,

samples have been post-fixed 1 hour at R.T. in 4% PFA+0.05% tween-20 in PBS 1x and then washed 3x15' in PBT. Embryos have been placed 10' in the pre-hybridization solution (50% formamide, 6x SSC, 0.05% tween 20) and finally 2 hour at 55°C in hybridization solution (50% formamide, 1x Denhardt's solution, 6xSSC, 0.05% tween 20, 100 µg/ml Yeast tRNA, 0,005% Heparine). Riboprobe (or Riboprobes for the double *in situ* hybridization) has been added up to a final concentration of 0.5 ng/µl and the hybridization occurred over night at the same temperature.

Day 2. A series of washes have been carried out by varying the temperature and salinity conditions; embryos have been, in fact, washed at 55°C in the following solutions: two times for 20' in Washing Buffer 1 (WB1: 50% formamide, 5x SSC, 0,1% SDS), two times for 20' in WB1:WB2 and two times for 20' in WB2 (50% formamide, 2x SSC, 0,1% Tween 20). Subsequently they have been treated with a 1 ml Solution A (10mM Tris-Cl, pH8.0, 0.5M NaCl, 5mM EDTA, 0.1% tween), two times for 5' at R.T. To remove aspecific RNA, not bound to the corresponding endogenous mRNA, the embryos have been treated with RNase A (20 µg/ml) for 20' at 37 °C in solution A, then washed one time in WB3 (2x SSC, 0.1% Tween20) for 5' at R.T. and 2 times in WB3 at a temperature of 55°C. Following they have been incubated three times for 5' in TNT (0.1M Tris, pH 7.5, 150mM NaCl, 0.1% tween) at R.T.

At this point a slight difference in protocol for single or double *in situ* occurs.

For single *in situ* hybridizations, embryos have been incubated in Blocking TNB buffer (100mM Tris pH7, 150mM NaCl, 1% Blocking Reagent, 0.2% Triton-100x) for 2 hours at R.T. At this point embryos have been incubated, all night at 4°C, with the antibody anti-DIG in the ratio 1:2000 in Blocking TNB Buffer.

For double *in situ* hybridizations, embryos have been first incubated in TNB blocking buffer for 2 hours at RT too. At this point samples have been incubated over night at 4°C, with anti-DIG Fab Fragments POD HRP diluted 1:400 in Blocking Buffer TNB.

Day 3. For single *in situ* hybridizations, the samples have been washed at R.T. in TNT with the following modalities: one time for 5', four times for 20', one time for 40' and three times for 10'

in TMN (100mM NaCl, 50mM MgCl2, 100mM Tris-Cl, pH9.5, 0.1% Tween20). To identify the localization of the RNA of interest, labeled with DIG and recognized by anti-DIG alkaline phosphatase conjugated, are provided the appropriate substrates that will be converted by the phosphatase in a precipitate of blue.

The embryos are incubated, therefore, in 1 ml of TMN containing 4.5 µl of NBT (nitroblue tetrazolium) and 3.5 µl of BCIP (5-bromo-4-chloro-3-indolyl-phosphate). The time of formation of the precipitate are conditioned by the amount of antibody bound and, therefore, indirectly on the type of probe used. For this reason, at intervals of 30' a few embryos are taken and observed, after being placed on a microscope slide, with a phase contrast microscope. When some signal was shown, the color reaction is stopped using 1x PBT.

Also for double *in situ* hybridizations series of washes were carried out: four times for 15' and two times for 5' in TNT at R.T. To identify the localization of the RNA of interest marked with Digoxigenin and recognized by anti-DIG conjugated to alkaline peroxidase (HRP) a substrate is converted by HRP in fluorescent product. The embryos have been incubated, therefore, in the 1x Plus Amplification Diluent (Perkin-Elmer) for 1' at R.T. and subsequently 1:400 Cy3 diluted in the same solution for 1.5 hours at R.T.

After the reaction, embryos were washed nine times for 5' in TNT at RT. In order to stop the first antibody reaction an incubation in 50% formammide, 2x SSC, 0,1% tween-20 for 10 minutes at 55°C has been performed. The next step consisted in a series of three washes of 10' in TNT and incubates, overnight at 4°C, with anti-Fluorescein HRP diluted 1:400 in Blocking Buffer TNB.

Day 4 (only for double *in situ*). The embryos have been washed two times for 5' in TNT at R.T., incubated in the 1x Plus Amplification Diluent (Perkin-Elmer) for 15'. and subsequently treated with 1:400 Cy5 diluted in the same solution for 1.5 hours. After this time the reaction has been blocked by carrying out the following washings: 1 time for 5' and 7 times for 10' in TNT at R.T. When desired, embryos were treated three times for 5 minutes and then over night with 1:10000 DAPI in TNT, in order to highlight the nuclei of the cells.

Single or double *in situ* hybridization experiments have been conducted at least three times independently, using a minimum of 25 embryos in each case. WMISH conducted on electroporated embryos were considered reliable when consistent in at least 60% of the analyzed embryos.

2.20 Preparation of constructs

Different constructs were prepared during this PhD study, the cloning strategies of them are explained below.

- *pGsx>mCherry*. *Ci-gsx* regulatory region, *pGsx*, was obtained by PCR amplification on genomic DNA using different combinations of specific primers designed according to the sequence of *C. intestinalis* genome (http://genome.jgi-psf.org/ciona4/ciona4.home.html). The amplification was done for a region of almost 3 kb upstream from Ci-gsx translation start site. Using the couple of oligonucleotides *pGsx1F - pGsx5R* I obtained a specific PCR product, of the expected length, that was cloned in in pCR®II vector (TOPO® TA Cloning Dual Promoter Kit, Invitrogen), following the manufacturer's indications.

  The pCR®II vector containing a *pGsx* fragment suitably orientated for the subsequent steps have been selected by PCR screening (paragraph 2.13), using the primers' couple *pGsxIntF* (designed within the insert, about 200 bp upstream from the 3' end of the fragment) and *M13Fwd* (it anneals in pCR ®II vector, downstream from the cloning site).

  The selected vector has been double digested with HindIII/NotI endonucleases (Roche), in order to excise a sticky-ended restriction fragment, *pGsx*, now suitable for the following cloning in the reporter containing plasmid. This fragment has been indeed purified by gel extraction kit (paragraph 2.7) and subsequently cloned upstream from mCherry reporter coding sequence, in place of *pTyrp1/2a* in *pTyrp1/2a>mCherry* vector (Esposito *et al.*, 2012), previously digested with HindIII/NotI (Roche) to eliminate *pTyrp1/2a*. The ligation reaction has been conducted as described in paragraph 2.11.

In the ***table 2.2*** are listed the oligonucleotide sequences used for pGsx amplification (in blue) and for the PCR screening (in black).

Table 2.2 Oligonucleotides used for *pGsx* cloning

Oligonucleotide name	Oligonucleotide sequence
pGsx1F	5'-CCGTCCTGTCTTACCCTGTGTGCTTTGAAA-3'
pGsx2F	5'-GTGCGAATCGCTAAACACTCTCTCGCGTGT-3'
pGsx3F	5'-GGTATGATAGTACCAGTTCGTGGTATCACA-3'
pGsx4R	5'-GGTTGTTGCTCAAGATCTATAAGAAAACGA-3'
pGsx5R	5'-TAGCCGCTTCAGTTCCTTATATAGCTATTAG-3'
pGsxIntFwd	5'-GAAGTACAGAGTATACCACAGATG-3'
M13Fwd	5'-CAGGAAACAGCTATGAC-3'

❖ *pGsx dissection*. As described in the paragraph 1.3, the comparison between *C. intestinalis* and *C. savignyi* regulatory region allows the efficient identification of conserved regulatory sequence information (Bertrand *et al.*, 2003; Johnson *et al.*, 2004; Squarzoni *et al.*, 2011). The alignment of the 5' upstream region of *Ci-gsx* with the corresponding region of *C. savignyi* revealed three regions that are conserved between the two species, and I named them I, II, III (***figure 2.1 A***, in pink). Specific oligonucleotide have been designed in order to test different combinations of conserved/non-conserved regions cloned upstream from the reporter mCherry (***figure 2.1***), later tested in transgenesis *via* electroporation experiments. The oligonucleotide sequences include the recognition site for restriction enzymes suitable for the subsequent cloning (***table 2.3***; HindIII site is indicated in red; EcoRV site are underlined; SmaI restriction site is indicated in light green). The oligonucleotide TATAGsx4R (indicated as 4' in table 2.3) also contains the sequence of the Epstein Barr virus TATA (E1bTATA), functioning in *Ciona* as minimal promoter (Squarzoni *et al.*, 2011; ***figure 2.1 B***, purple asterisks, and ***table 2.3***, purple sequence). All the fragments, except *pGsxD*, have been inserted in place of *pGsx* in *pGsx>mCherry* vector, previously digested with HindIII/EcoRV enzymes to remove *pGsx*.

The PCR products *pGsxA*, *pGsxB* and *pGsxI*, obtained using the combination of oligonucleotides represented in ***figure 2.1 B***, have been digested HindIII/EcoRV and

cloned upstream from mCherry reporter as previously described, in order to obtain *pGsxA>mCherry*, *pGsxB>mCherry* and *pGsxI>mCherry*.

To obtain *pGsxC>mCherry* and *pGsxIII>mCherry*, *pGsxC* and *pGsxIII* PCR products have been digested with EcoRV enzyme, and then cloned upstream from mCherry, in a HindIII/EcoRV digested vector, in which the 5' overhang released by HindIII cleavage was previously filled in as described in the paragraph 2.9.

Multiple ligase reactions have been set up in order to clone two or three different PCR fragments, conveniently digested, in the mCherry containing vector (*pGsxE>mCherry*, *pGsxF>mCherry*, *pGsxU>mCherry*, *pGsxI+III>mCherry*). The protocol is similar to the one reported in paragraph 2.11; the proportion of plasmid vector and each DNA insert was kept 1:3. PCR screening and careful sequence analyses have been done to check the proper orientation of each fragment.

In order to obtain *pGsxD*>mCherry construct, *pGsxD* PCR fragment has been double digested with HindIII/EcoRV endonucleases, and then cloned in place of *pGsxI* in *pGsxI>mCherry* plasmid, previously digested HindIII/SmaI to excise pGsxI, in order to take advantage of the E1bTATA sequence. Indeed, both EcoRV and SmaI enzymes release blunt ends after restriction, and their restriction products are therefore compatible.

Table 2.3 Oligonucleotides used for *pGsx* dissection

Oligonucleotide name	Oligonucleotide sequence
1)HindIIIGsx1F	5'-CTGaagcttACTGAGCACTTAGTGCG-3'
2)EcoVGsx2F	5'-TATgatatcGGATGCTCTAAATGTAC-3'
3)EcoVGsx3F	5'-TATgatatcAGATCCTAAACCGACTAC-3'
4)EcoVGsx4R	5'-ATTgatatcGTCAGATTTCAACATGG-3'
4')TATAGsx4R	5'-ATTgatatcCATTATATAcccgggTCAGATTTCAACATGG-3'
5)EcoVGsx5R	5'-ATTgatatcGATGTGCGTGTACTTG-3'
6)EcoVGsx6R	5'-ATTgatatcAATGAACCTGCCTTTG-3'
7)EcoVGsx7R	5'-ATTgatatcTAGCCGCTTCAGTTCCTTATATAG-3'

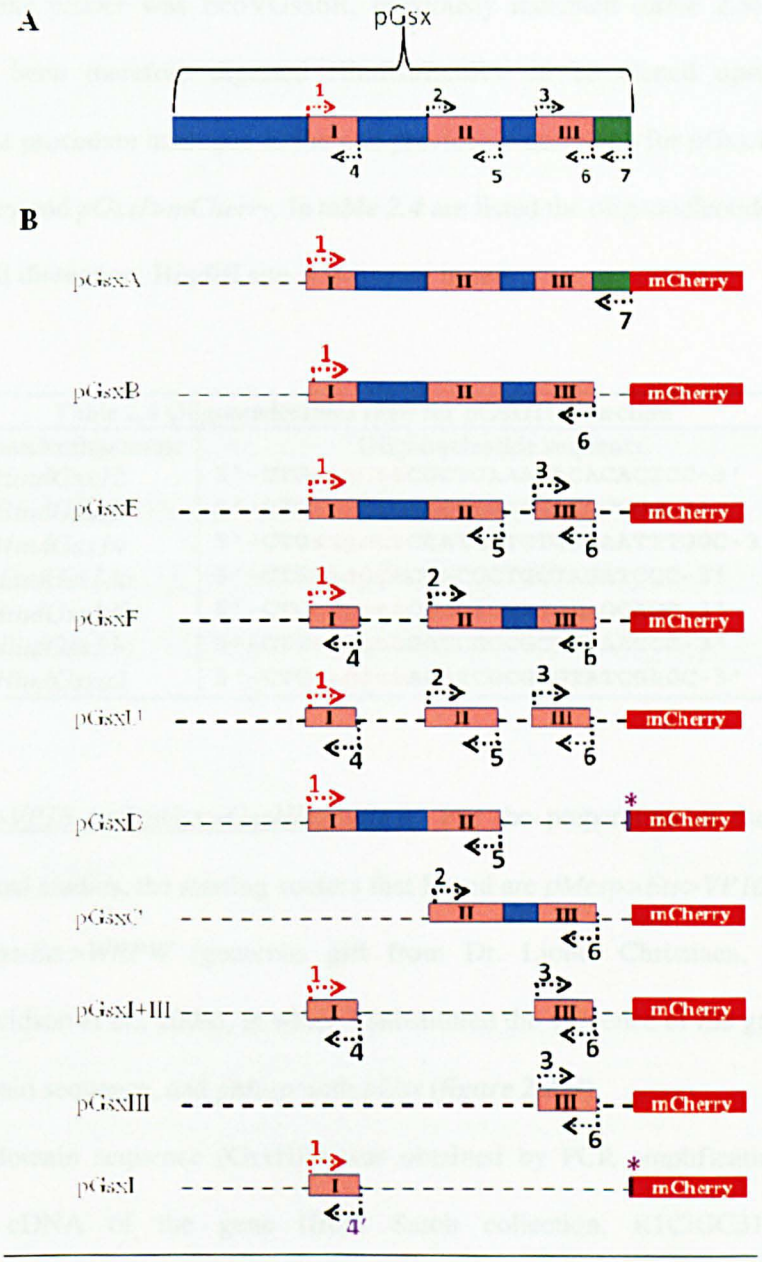


Figure 2.1 Deletion of *pGsx* regulatory region. **A.** Schematic representation of *pGsx* fragment. The pink boxes indicate the three conserved regions between *C. intestinalis* and *C. savignyi*. The light blue boxes indicate the non conserved regions. The green box represents part of *Ci-gsx* 5' UTR included in *pGsx* fragment. The arrows, with the respective numbers, indicate the position of the different oligonucleotides used for the promoter dissection. **B.** Schematic representation of the different constructs obtained in this study. The combination of primers used are indicated. Purple asterisks indicate the E1bTATA minimal promoter.

❖ *pGsxIII dissection*. A 5' deletion study of *pGsxIII* promoter region has been performed using a series of oligonucleotides specifically designed in order to amplify progressively shorter fragments. All the forward primers had the HindIII site at their 5' end. In all the

cases, the reverse primer was EcoVGsx6R, previously indicated (***table 2.3***). The PCR products have been therefore digested HindIII/EcoRV to be cloned upstream from mCherry, with a procedure analogue to the one previously described for *pGsxA>mCherry*, *pGsxB>mCherry* and *pGsxI>mCherry*. In ***table 2.4*** are listed the oligonucleotide sequences used for pGsxIII dissection; HindIII site is indicated in red.

Table 2.4 Oligonucleotides used for pGsxIII dissection

Oligonucleotide name	Oligonucleotide sequence
HindGsx12	5'-CTGaagcttCGCTGAAACACACACTCC-3'
HindGsx13	5'-CTGaagcttAAGCCGGCGATCCGGG-3'
HindGsx14	5'-CTGaagcttCCATGATCTGGTAATTTGGC-3'
HindGsx13a	5'-CTGaagcttCGGCGCTGCTAGATCGC-3'
HindGsx13b	5'-CTGaagcttCGCGAGTATCGAGCAGC-3'
HindGsx13c	5'-CTGaagcttGATCGCCGCTTAAACCG-3'
HindGsxa2	5'-CTGaagcttAGATCGCGAGTATCGAGC-3'

❖ *pGsx>GsxHD>VP16 and pGsx>GsxHD>WRPW*. For the preparation of the constructs used in functional studies, the starting vectors that I used are *pMesp>Ets>VP16* (***figure 2.2 A***) and *pMesp>Ets>WRPW* (generous gift from Dr. Lionel Christiaen, New York University; Davidson *et al.*, 2006), in which I substituted the sequence of *Ets* gene with *Ci-gsx* homeodomain sequence, and *pMesp* with *pGsx* (***figure 2.2 B***).

Ci-gsx homeodomain sequence (GsxHD) was obtained by PCR amplification using as template the cDNA of the gene (from Satoh collection, R1CiGC31m18). The oligonucleotides used (***table 2.5***) contained NheI at the 5' end and SpeI at the 3' end. The digested PCR product was ligated in place of Ets in *pMesp>Ets>VP16* and *pMesp>Ets>WRPW* vectors, in wich Ets was previously excised thanks to NheI/SpeI restriction (***figure 2.2 A***).

Subsequently, *pMesp* has been removed by XhoI/Not digestion, and replaced with pGsx, obtained from *pGsx>mCherry* previously digested with the same enzymes (***figure 2.2 A***).

In ***table 2.5*** are listed the oligonucleotide used for GsxHD amplification; NheI restriction site is indicated in light blue, SpeI site in brown.

Table 2.5 Oligonucleotides used for GsxHD amplification	
Oligonucleotide name	**Oligonucleotide sequence**
NheGsxHDf	5'-ATgctagcTCAAGTTAGCAGTCGGGTAACA-3'
SpeGsxHDr	5'-ATactagtAGTACACTGTTCCCGCGAGTCCTGCT-3'

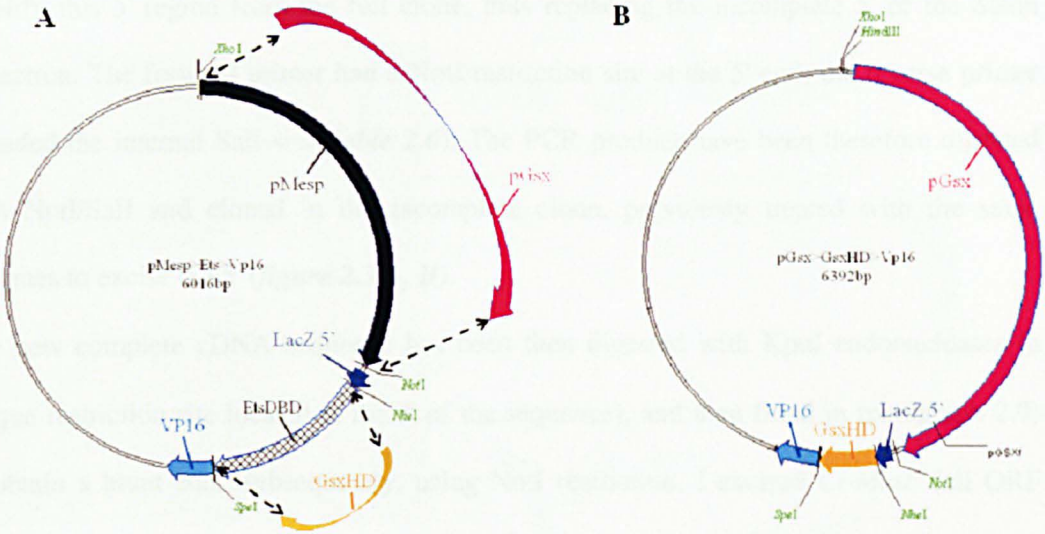


Figure 2.2 Preparation of *pGsx>GsxHD>VP16* construct. Restriction sites are indicated in green; the Ets Dna binding Domain (DBD) is in frame with the 5' of LacZ. pMesp has been substituted with pGsx (XhoI/NotI restriction) and EtsDBD with GsxHD (NheI/SpeI restriction). See text for details. The strategy to obtain *pGsx>GsxHD>WRPW* from *pMesp>Ets>WRPW* is the same.

- *pGsx>LacZ*. pBlueScript 1230 (gift of R. Krumlauf, Stowers Institute, Kansas City, USA), which contains the LacZ reporter gene and SV40 polyadenylation sequence with the human β-globin basal promoter, was used to generate *pGsx>LacZ* construct, often used as control plasmid in my experiments. *pGsx* was excised from *pGsx>GsxHD>VP16* plasmid with KpnI digestion, and then cloned upstream from LacZ reporter gene, in pBlueScript 1230 previously digested with the same enzyme. The correct 5'-3' orientation of the promoter was checked by PCR screening (paragraph 2.13).
- *pGsx>Msx*. *Ci-msxb* cDNA sequence present in N. Satoh *C. intestinalis* gene collection 1 (R1CiGC42h24) was lacking of the 5' (***figure 2.3 B***). Besides this gene collection, it is also available in the laboratory a Gateway full ORF library, in which the full coding sequence

of *Ciona* genes are present; here, the corresponding clone of *Ci-msxb* is cien92596 (plate: VES84_L19; ***figure 2.3 A***). Unfortunately, the polylinker of this vector is too limited for a further subcloning of the cDNA of interest. In order to avoid the PCR amplification of the full length gene, a procedure that could possibly insert mutations in the coding sequence, I decided to use a different approach. *Ci-msxb* sequence presents an internal SalI restriction site 452 bp from the starting ATG codon. I designed specific oligonucleotides in order to amplify this 5' region from the full clone, thus replacing the incomplete 5' of the Satoh collection. The forward primer had a NotI restriction site at the 5' end; the reverse primer included the internal SalI site (***table 2.6***). The PCR product have been therefore digested with NotI/SalI and cloned in the incomplete clone, previously treated with the same enzymes to excise the 5' (***figure 2.3 A, B***).

The new complete cDNA sequence has been then digested with KpnI endonucleases (a unique restriction site located at the 3' of the sequence), and then filled in (paragraph 2.9) to obtain a blunt end. Subsequently, using NotI restriction, I excised *Ci-Msxb* full ORF from this starting vector (***figure 2.3 C***).

In parallel *pGsx* sequence has been obtained from *pGsx>mCherry* vector, with XhoI/NotI restriction.

The final receiving vector was *pTyrp1/2a>FGFR^{DN}>mCherry*, available in the lab (Squarzoni *et al.*, 2011), in which the whole pTyrp1/2a>FGFR^{DN} sequence was removed by double restriction XhoI/PvuII (PvuII releases blunt ends after cleavage).

At this point a triple ligase reaction has been set up, using the two inserts *pGsx* (XhoI/NotI), Ci-*msxb* (NotI/filled in blunt end), and the vector digested with XhoI/PvuII enzymes. The protocol is similar to the one reported in paragraph 2.11 (***figure 2.3 D***); the proportion of plasmid vector and each DNA insert was kept 1:3. PCR screening and careful sequence analyses have been done to check the sequence and correct insertion of each fragment.

Table 2.6 Oligonucleotides used for Ci-msxb 5' amplification	
Oligonucleotide name	**Oligonucleotide sequence**
NotMsxF	5'-ATTgcggccgcATTGTAGAATATTGAAGAATA-3'
SalMsxR	5'-ATTgtcgacTCGTAACTACAATC-3'

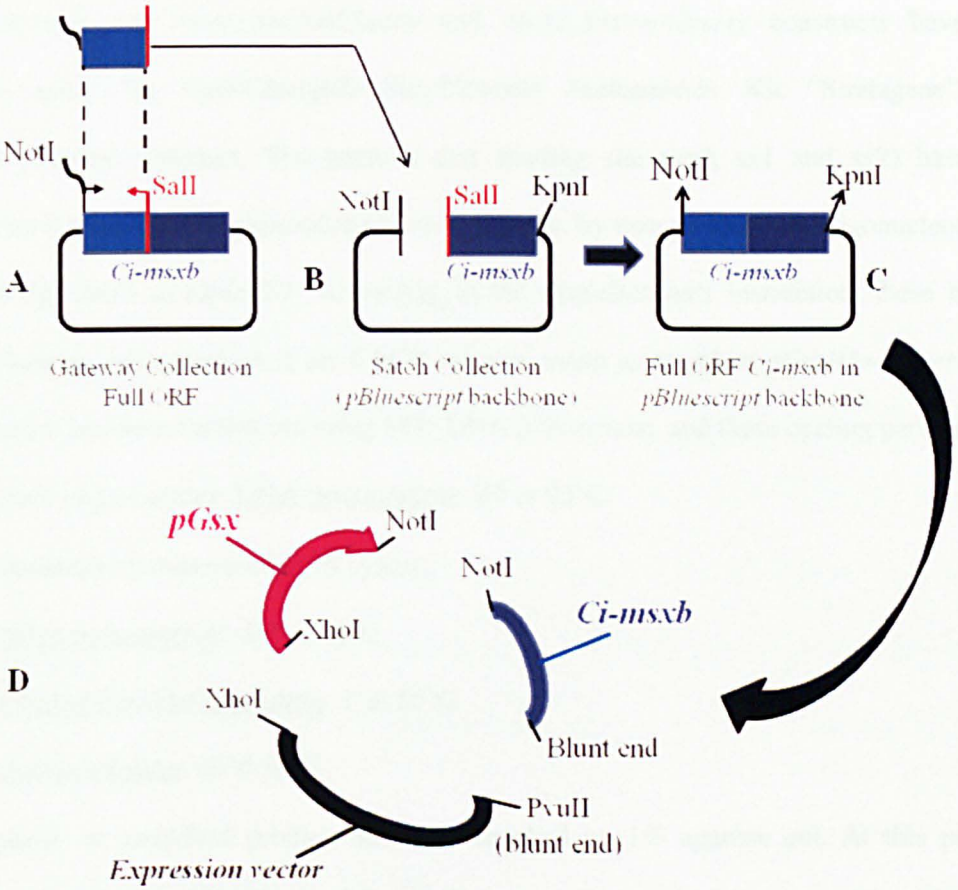


Figure 2.3 Preparation of *pGsx>Msx* construct. A. Schematic representation of *Ci-msxb* clone in the Gateway full ORF library (VES84_L19). In red, SalI restriction site, which is internal to *Ci-msxb* sequence, is represented. NotI restriction site was inserted by PCR. **B.** Schematic representation of Ci-msxb clone in Satoh collection, from which the incomplete 5' has been removed through NotI/SalI restriction. **C.** Schematic representation of the full ORF sequence of Ci-msxb obtained. D. Schematic view of the triple ligase reaction to obtain *pGsx>Msx* construct. *pGsx* (in pink) was obtained from pGsx>mCherry construct. See text for details.

2.21 *In silico* analysis of putative *trans*-acting factors

The pGsxIII sequence was submitted to JASPAR database (http://jaspar.genereg.net/) and the Genomatix Database (http://www.genomatix.de/cgi-bin/eldorado.main.pl), using in both cases the default parameters. These are databases of transcription factor from many organisms, their genomic binding sites and DNA-binding profiles. JASPAR analysis revealed the presence of

three putative Sox binding sites. Genomatix analyses revealed the presence of two putative Msx binding sites.

2.22 Site directed mutagenesis of putative Sox binding site on pGsxIII regulatory region

Mut0_Sox>mCherry, *Mut1_Sox>mCherry* and *Mut2_Sox>mCherry* constructs have been prepared using the QuikChange® Site-Directed Mutagenesis Kit "Stratagene" from *pGsxIII>mCherry* construct. The putative Sox binding site (sx0, sx1 and sx2) have been replaced by a sequence that reduced the binding affinity, by using mutagenic oligonucleotides of about 50 bp, listed in ***table 2.7***. According to the manufacturer's instruction, these mutated oligonucleotides have been used for a PCR reaction using as template *pGsxIII>mCherry*. The PCR reaction has been carried out using PFU DNA polymerase, and these cycling parameters:

- First step (1 cycle). DNA denaturation: 30" at 95°C.
- Second step (repeated for 18 cycles).

  DNA denaturation: 30" at 94°C.

  Oligonucleotides annealing: 1' at 55°C

  Polymerization: 68°C for 7'.

The presence of amplified product has been checked on 1% agarose gel. At this point, to eliminate the wild type plasmid, used as template, the mixture has been digested at 37°C for 1.5 hour, with 1 µl of the *Dpn*I restriction enzyme (10 U/µl). This enzyme digests only the methylated supercoiled dsDNA used as template, but not the newly synthesized mutated and unmethylated DNA. After the digestion, an aliquot of the reaction has been transformed and grown as previously described (paragraph 2.12). Ten clones have been selected after growth; the isolated plasmids have been sequenced to check the presence of mutations.

In ***table 2.7*** the mutagenic oligonucleotide sequences are reported. The wild type sequence for each putative Sox binding site is also indicated; putative Sox binding sites are underlined and the corresponding mutated sequence is indicated by small, underlined, red letters.

The mutagenesis for each site allowed me to obtain *Mut0_Sox>mCherry*, *Mut1_Sox>mCherry* and *Mut2_Sox>mCherry* constructs.

In order to obtain the double mutant for sx1 and sx2 binding sites (*DbM_Sox>mCherry*), I mutated sx2 site in the single mutant *Mut1_Sox>mCherry*.

The mutagenesis of sx0 site on *DbM_Sox>mCherry* construct allowed me to obtain the triple mutated plasmid, *TrM_Sox>mCherry*.

In order to obtain *pGsx12DM>mCherry* construct, I performed a PCR using *DbM_Sox>mCherry* as template, and the couple of primer *HindGsx12* (***table 2.4***) and EcoVGsx6R (***table 2.3***). The amplified product has been digested with HindIII/EcoRV restriction enzymes and cloned upstream from mCherry as previously described.

Table 2.7 Mutagenic oligonucleotides

Sx0: wild type sequence	CCGGTTAATTGTAACGGAGATTTAAACAAGGTCGCTGAAACACACACTCCC
Oligonucleotide name	**Oligonucleotide sequence**
Mut0SoxFwd	CCGGTTAATTGTAACGGAGATTTAccggggGTCGCTGAAACACACACTCCC
Mut0SoxRev	GGGAGTGTGTGTTTCAGCGACccccggTAAATCTCCGTTACAATTAACCGG
Sx1: wild type sequence	CCCTGGTCTCCAACTGTCATCAAACAATTAAGCCGGCGATCCGGGGTTC
Oligonucleotide name	**Oligonucleotide sequence**
Mut1SoxFwd	CCCTGGTCTCCAACTGTCATCAccCccgTAAGCCGGCGATCCGGGGTTC
Mut1SoxRev	GAACCCCGGATCGCCGGCTTAcggGggTGATGACAGTTGGAGACCAGGG
Sx2: wild type sequence	CTTATGATCGCCGCTTAAACCGATTTGTTTCCATGATCTGGTAATTTGGC
Oligonucleotide name	**Oligonucleotide sequence**
Mut2SoxFwd	CTTATGATCGCCGCTTAAACCGAaccccccCCATGATCTGGTAATTTGGC
Mut2SoxRev	GCCAAATTACCAGATCATGGcgggggggtCGGTTTAAGCGGCGATCATAAG

2.23 Quantitative Real-time PCR (qPCR)

2.23.1 RNA extraction and RNA quality detection

RNA extraction have been performed using the RNAqueous-micro kit (Ambion). Transgenic embryos at the desired developmental stage have been collected in a 1.5 ml eppendorf tube and pelletted by centrifugation for 5', 4000 rpm. The supernatant sea water was quickly removed and replaced with 250 µl of lysis solution, provided with the kit. The centrifugation step is useful to remove supernatant sea water from embryos at late tailbud stage, since they do not settle by their own, as occurs instead for previous developmental stage. The embryos in the lysis buffer have been stored at -80°C or immediately used for RNA extraction, performed according to manufacturer's recommendations. Total RNA extracted (1 µl) has been analyzed on the

BioAnalyzer machine using the Eukaryote total RNA Nano Series II kit (Agilent Technologies; ***figure 2.4***).

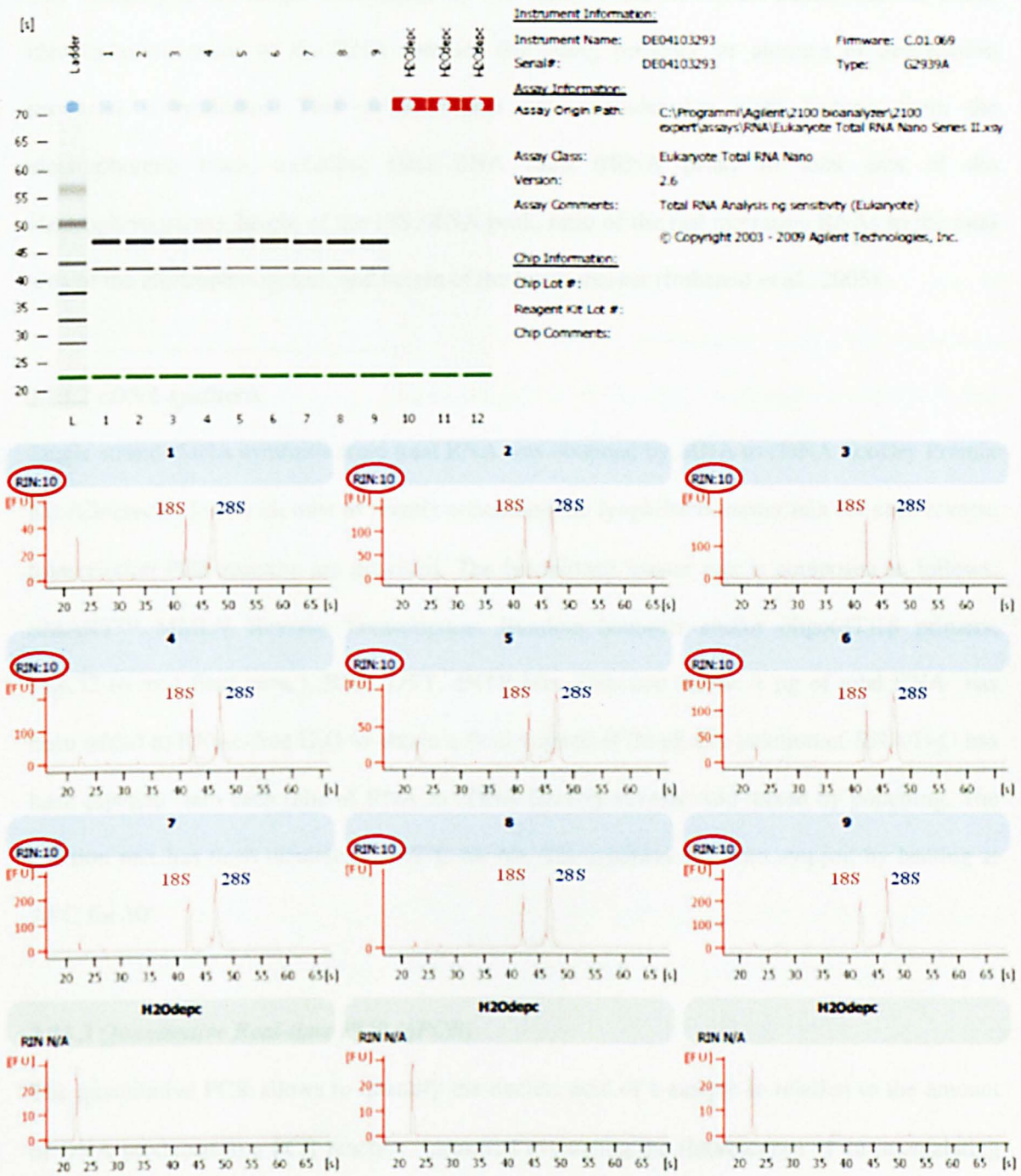


Figure 2.4 Summary of the electrophoresis profiles generated by the BioAnalyzer. The upper panel shows the virtual electrophoresis gel image of 9 RNA samples (lane 1-9), used in my experiments; electrophoretic profiles of the ladder (first lane) and blanks (lanes named H20depc) are also present.
The lower panels show the electropherogram profiles summary of my RNA samples (1-9) and blanks (H20depc). The RIN score for each sample is highlighted by red circles.

The integrity of total RNA is estimated by RNA Integity Number (RIN) values, calculated by an algorithm that assigns a 1 (completely degraded RNA) to 10 (intact RNA) RIN score. In this way integrity is no longer determined by the ratio of the ribosomal bands, but the entire electrophoretic trace of the RNA samples (including presence or absence of degradation products) is evaluated. This method takes into consideration eight features from the electrophoretic trace, including total RNA ratio (rRNA peaks to total area of the electropherograms), height of the 18S rRNA peak, ratio of the fast migrating RNAs to the total area of the electropherograms, and height of the lower marker (Imbeaud *et al.*, 2005).

2.23.2 cDNA synthesis

Single strand cDNA synthesis from total RNA was obtained by RNA to cDNA EcoDry Premix Kit (Clontech). In this kit tube of premix containing the lyophilized master mix for each reverse transcription PCR reaction are provided. The lyophilized master mix is composed as follows: SMART™ MMLV Reverse Transcriptase, Random hexamer and/or oligo(dT)18 primers, MgCl2 (6 mM final conc.), BSA, DTT, dNTP Mix, Reaction buffer. 1 µg of total RNA has been added to RNase-free H_2O to obtain a final volume of 20 µl; this solution of RNA/H_2O has been pipetted into each tube of RNA to cDNA EcoDry Premix, and mixed by popetting. The reaction mix has been incubated at 42°C for 60'. The reaction has been stopped by heating at 70°C for 10'.

2.23.3 Quantitative Real-time PCR (qPCR)

The quantitative PCR allows to quantify the nucleic acid of a sample in relation to the amount of DNA produced in a PCR reaction, measured evaluating the fluorescence of an intercalating DNA dye (in the case of the approach that I used), which is monitored at each cycle during the amplification. This is proportional to the amount of amplicon production. During the PCR exponential phase the amount of product increases linearly (on a log plot). Therefore, PCRs from different samples can be compared to determine which among them contains the higher amount of a specific sequence. This is done by setting a threshold in the linear phase of

amplification and assessing how many cycles (Ct) are necessary to reach the same level of product amplification in different samples.

Because of the chemistry used, when the difference in the amount of a specific sequence between two samples (a "control" and an "experimental" one) is assessed, this is given by $2^{\Delta Ct}$, where 2 is the multiplier for amplification per cycle, and ΔCt is the difference in the Ct between the two samples. The dye used was Fast Sybr Green Master Mix (Applied Biosystems), which binds to double stranded DNA (dsDNA). SYBR green dye can not distinguish between the amplicon and contamination products from mispairing or primer-dimer artifacts. To overcome this, not only DNA synthesis is monitored, but also the melting point of the PCR products is measured at the end of the amplification reaction. The melting temperature of a DNA double helix depends on its base composition and its length. Therefore, qPCR products can be distinguished from mispairing or primer-dimer artifacts.

This method requires a known reference gene as internal control with constant expression in all tested samples and whose expression is not changed by the treatment under study. A pair of primers for Rps27 ribosomal gene, which is therefore ubiquitary in *Ciona* embryos, has been used as standard reference. In addition, I also used Brachyury (Bra) as reference gene, which is specifically expressed in Ciona notochord, in a limited number of cells, comparable to the number of cells of Ci-gsx lineage, in which I was interested in my experiments

For each gene qPCR primers have been designed to generate products of 100-200 bp, by using online based "Primer 3, v.0.4.0" software (http://fokker.wi.mit.edu/primer3/input.htm; **table 2.8**). Blast searches against the whole *Ciona* genome have been performed to verify primers specificity.

The number of cycles needed for the standards to reach a specified Ct is used to normalize the Ct for the selected genes. To capture intra-assay variability all RT-qPCR reactions have been carried out in triplicate and the average Ct value was taken in to account for further calculations. The efficiency of each pair of primers was calculated according to standard method curves using the equation E=10-1/slope. Five serial dilutions have been set up to determine the Ct value and the efficiency of reaction of all pairs of primers. Standard curves were generated for each

oligonucleotides pair using the Ct value versus the logarithm of each dilution factor. Diluted cDNA was used as template in a reaction containing a final concentration of 0.7 pmol/µl for each primer and 1X Fast SYBR Green master mix (total volume of 10 µl).

PCR amplifications have been performed in triplicate in a ViiA7 ABI Applied Biosystems thermal cycler, using the following thermal profile: 95°C for 20", one cycle for cDNA denaturation; 95°C for 1" and 60°C for 20", 40 cycles for amplification; 95°C for 15", 60°C for 1' and 95°C for 15", one cycle for melting curve analysis, to verify the presence of a single product. Each assay included a no-template control for each primer pair.

Real time PCR has been carried out in order to analyze expression changes between embryos with interfering Ci-gsx constructs (*pGsx>GsxHD>VP16* and *pGsx>GsxHD>VP16*) versus control embryos (electroporated with *pGsx>mCherry*). To determine the expression ratio between the sample and control embryos, Pfaffl Method based on this equation has been used (Fleige *et al.*, 2006):

$$\text{Ratio} = \frac{E\,(\text{Target})^{\Delta Ct\ \text{target(Control-Treated)}}}{E\,(\text{Reference})^{\Delta Ct\ \text{reference(Control-Treated)}}}$$

In the above equation, E target and E reference are the amplification efficiencies of the target and reference genes, respectively. ΔCt, target (Control–Treated) = Ct of the target gene in the control sample minus the Ct of the target gene in the treated sample, and ΔCt, reference (Control–Treated) is the Ct of the reference gene in the control sample minus the Ct of the reference gene in the treated sample. Data for each gene were normalized against Rps27 and Bra.

Table 2.8 Oligonucleotides used for qPCR

Oligo name	Oligonucleotide sequence
Rps27F	5'-AATCCACTCTTCACCTTGTG-3'
Rps27F	5'-GGGAGATCTTGCCATTTTC-3'
BraF	5'-CGCATTCATTTCAGGAGACA-3'
BraR	5'-TCGAATAATTTTGCGACGAA-3'
RxF	5'-CCTCGAACGAAGATCAAAGC-3'
RxR	5'-ACATCCGGGTAGTGAGATCG-3'
ArrF	5'-CAAGTGTCCTGTTGCCTTCA-3'
ArrR	5'-CCAGACCTCGCTTATCCTTG-3'
MsxF	5'-acgtcaaaggaatcgtcacc-3'
MsxR	5'-AGTAGTTGCTCGGTGCTGAA-3'

2.24 Alignments of Gsx sequences

Gsx protein sequences from *Mus musculus*, *Homo sapiens*, *Xenopus tropicalis*, *Danio rerio* *Branchiostoma floridae*, *Ptychodera flava*, *Patiria miniata*, *Platynereis dumerilii*, *Drosophila melanogaster*, *Anopheles gambia*, *Tribolium castaneum*, *Hydra vulgaris*, *Clytia hemisphaerica* and *Trichoplax adhaerens* were obtained using NCBI Protein Database (http://www.ncbi.nlm.nih.gov/protein/), and used in multiple alignment with Ci-gsx sequence. This multiple sequence alignment was done with Clustal Omega (CLUSTAL-O) software, specific for proteins (http://www.ebi.ac.uk/Tools/msa/clustalo/ ; Sievers *et al.*, 2011). Accession numbers for Gsx proteins used in these analyses are listed in ***table 2.9***.

Table 2.9 Accession Numbers of Gsx protein sequences used in the multiple alignment

Species	Name	Accession Number
Homo sapiens	Gsx1	EAX08418.1
	Gsx2	NP_573574.1
Mus musculus	Gsx1	NP_032204.1
	Gsx2	NP_573555.1
Xenopus tropicalis	Gsx1	NP_001039254.1
	Gsx2	AAI58504.1
Danio rerio	Gsx1	CAM14173.1
	Gsx2	AAI64330.1
Branchiostoma floridae	AmphiGsx	AC129948.3
Ptychodera flava	PfGsx	AAR07642.1
Patiria miniata	PmGsx	AGK89736.1
Platynereis dumerilii	PdGsx	ACH87538.1
Drosophila melanogaster	Dm-ind	AAS65004.2
Anopheles gambia	Ag-ind	CAD27926.1
Tribolium castaneum	Tc-ind	AAW21974.1
Hydra vulgaris	Cnox2	AAA29210.1
Clytia hemisphaerica	Cnox2	ACM62729.1
Trichoplax adhaerens	Trox2	AAS54997.1

Chapter 3

RESULTS

Ci-GSX PROTEIN SEQUENCE

3.1 Comparative analysis of Ci-gsx protein sequence

Gsx is a homeobox transcription factor whose presence has been detected (or at least predicted) in almost all the *taxa* of animal kingdom, with the exception of sponges. Vertebrate genomes contain two members of this family, while in invertebrates one member is reported. *Drosophila* Ind deduced protein, the Gsx family member best studied so far, contains four specific domains. The Eh1-related domain (Engrailed homology 1), at the N-terminal, which is involved in Ind transcriptional repression activity through the interaction with the co-repressor Groucho; the Pst domain, at the N-terminal too, which has transcriptional repression ability that appears to act independent of interaction with the co-repressor Groucho; the homeodomain, at the C-terminal, responsible for the interaction with DNA; an activation domain, between the Pst domain and the homeodomain (Von Ohlen and Moses, 2009; Von Ohlen *et al.*, 2007a).

As previously mentioned, a single ortholog of *Gsx ParaHox* gene family is present in *Ciona intestinalis* genome, called *Ci-gsx*. In order to verify the presence of any evolutionary conserved domain (besides the homeodomain) in Ci-gsx, its deduced amino acid sequence was analyzed, in a comparative study with Gsx homologs present in vertebrates (*Mus musculus, Homo sapiens, Xenopus tropicalis, Danio rerio*), amphioxus (*Branchiostoma floridae*), in the hemichordate *Ptychodera flava*, in the sea star *Patiria miniata*, in the annelid *Platynereis dumerilii*, in insects (*Drosophila melanogaster, Anopheles gambia, Tribolium castaneum*), cnidarians (*Hydra vulgaris* and *Clytia hemisphaerica*) and in the placozoan *Trichoplax adhaerens*.

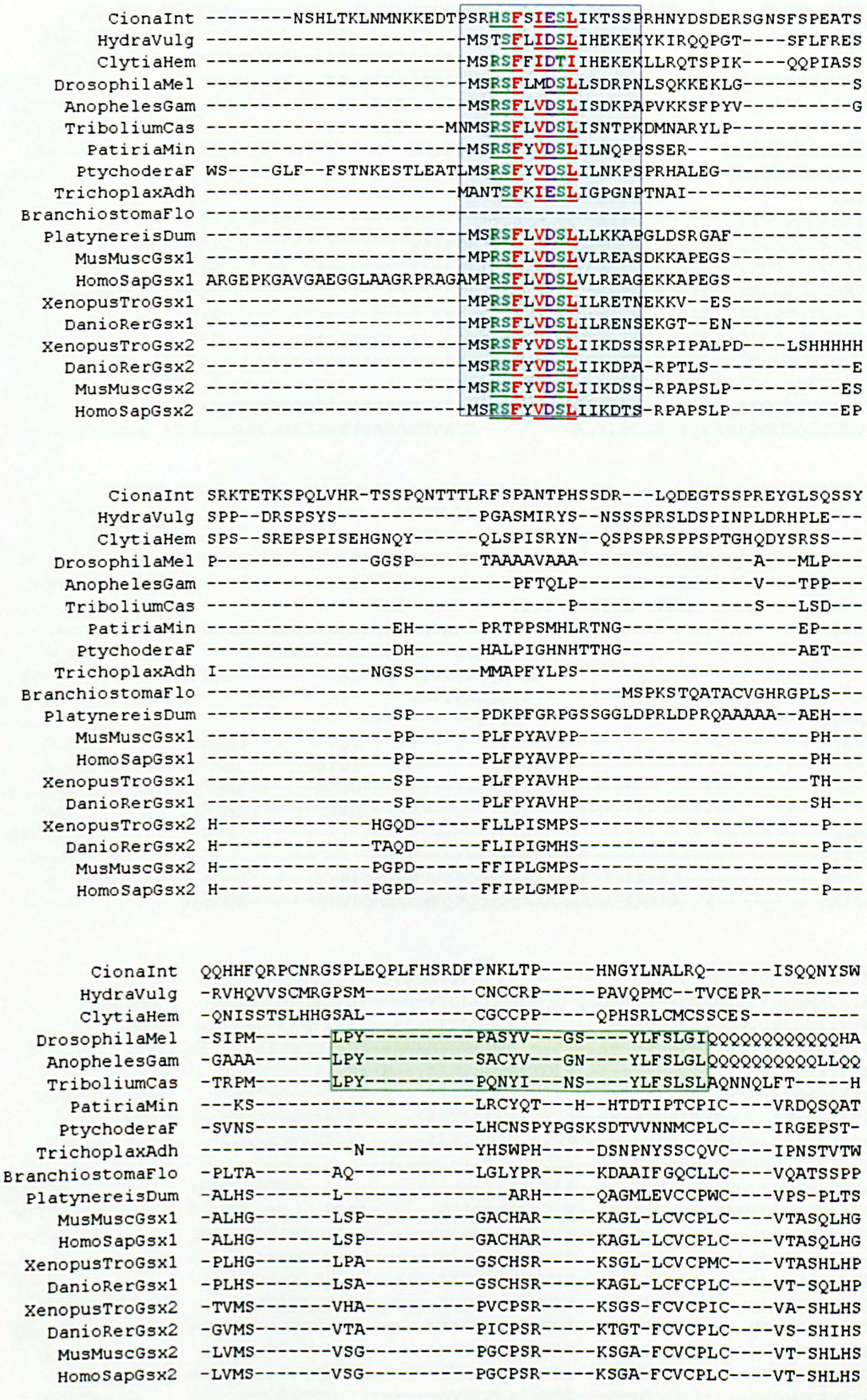
The ***figure 3.1*** reports the results of this analysis. Eh1-related domain is highlighted by a blue box and the homeodomain by the pink box. The two insect-specific regions, PstI (green box) and *Drosophila* activation domain (yellow box) are also reported.

The alignment done with CLUSTAL-O permitted to verify that the conservation was noticeable in the homeodomain amongst all the species examined, but this is not a novelty. Furthermore, a good grade of conservation was present at the level of the Eh1 domain in all the organism included, except amphioxus.

The consensus motif that is described as typical "Engrailed homology domain" is constituted by three conserved amino acids: phenylalanine-x-isoleucine-x-x-isoleucine (F-x-I-x-x-I), where x is any amino acid. However, it has been already reported that in most cases isoleucine can be replaced by valine (V) or leucine (L) (Goldestein *et al.*, 2005). In addition, it is also known that the presence of negatively charged amino acids at positions 4 or 5 (between the two Ile) are quite common, as well as the presence of potential phosphoacceptor serines, in position 2, adjacent to the invariant Phe residue (Goldestein *et al.*, 2005).

The alignment clearly indicated that in all the Gsx sequences analyzed, including Ci-gsx, the main features of the general Eh1 domain were conserved (red and purple amino acids in **figure 3.1**). Among the analyzed proteins, only Ci-gsx presented a serine in position 2. Putative phospoacceptor serines are actually present in the aligned sequences, but they lie in position -1 and 5 with respect to the conserved Phe. In addition, I also observed the presence of positively charged amino acids in position -2. These features that were not previously described for the general Eh-1domain are represented by green residues in ***figure 3.1***.

The analysis confirmed that the Pst domain (green box in the figure) is a peculiarity of insect Gsx proteins (Von Ohlen, 2009).


CionaInt
HydraVulg
ClytiaHem
DrosophilaMel
AnophelesGam
TriboliumCas
PatiriaMin
PtychoderaF
TrichoplaxAdh
BranchiostomaFlo
PlatynereisDum
MusMuscGsx1
HomoSapGsx1
XenopusTroGsx1
DanioRerGsx1
XenopusTroGsx2
DanioRerGsx2
MusMuscGsx2
HomoSapGsx2

```
          CionaInt NDRAAA-------------ALVTFS--------VLGGNFGQWQN-------NNLNYPHCS
         HydraVulg -----EPGEGT---SS-----------------------Q------Y----------PY-
         ClytiaHem -----GPREGE---AA-----------------------P------F----------GS-
     DrosophilaMel AAAAAAAA------AA--AALQQHPHVSSSPGSLYHPYAQLFA-SKRKSS-GFSNYEGCY
      AnophelesGam HPPSVFP--------H--AAL---PHTPT-----NQPIVPFWQ-ADLKPELSPKLYQPTD
      TriboliumCas HKAPIRP--------V--ALP---P-----------------------------------
        PatiriaMin CVGRLP---------------GMPPPT--------SAAVAVFK-----SPLASLAPSHHS
       PtychoderaF CST-IN---------------NM---V--------TTTVSMFK-----PPSSVVNVMSH-
      TrichoplaxAdh PVLVK-----------------------------PSPIVD--------------PAI
  BranchiostomaFlo ---TT--GHSHF-------------------------GIPRTS-----VPTTVPSFASTL
    PlatynereisDum LNKSTTSA-------------STIPHH--------HAPHPIFSMAPLHPPPGAPPVAPRP
       MusMuscGsx1 PPGPP-----------------------------ALPLLK-ASF--PPFGSQYCHAP
       HomoSapGsx1 PPGPP-----------------------------ALPLLK-ASF--PPFGSQYCHAP
     XenopusTroGsx1 --PPP-----------------------------GIPLLK-ASF--SSFGTQYCPAG
       DanioRerGsx1 --SPP-----------------------------ALPLIK-ASF--PAFGTQYCHSA
     XenopusTroGsx2 SRNSI---------------------------PLLK-GHF--PNGEAQYCQRV
       DanioRerGsx2 PRGGI---------------------------PLLK-GQ-GFTAGDAAFCQRM
       MusMuscGsx2 SRPPAGAGGGATGTAGAAVAGGGVAGG--------TGALPLLK-SQFSPAPGDAQFCPRV
       HomoSapGsx2 SRGSVGAGSGGAGAGVTGAGGSGVAGA--------AGALPLLK-SQFSSAPGDAQFCPRV

          CionaInt ERDIQDFQASSYSHAPDLRSIHRNPAYHRNDPS---YM----------------VDQSQ
         HydraVulg ------------TREP---HEHTRGLY-GNDRSRL-FPIL--------------------
         ClytiaHem ------------PREA--THHTTRYLYGGSERGRI-FSLA--------------------
     DrosophilaMel PSPPLSA-NPNSQQLPPIHNLYGSPVVGGLPLPEPGSFCTSPSASSSASLDYTNNFDEPQ
      AnophelesGam VSP---------PKLSPKEDYYR-KLYRGEKTPERYAPYL--AVRPVESLTSGSNVAAAA
      TriboliumCas -----------------KRE---------IPSPKRSSPIL--AEK---------------
        PatiriaMin PSEHHRM-LPTRSYLESCRDHHLHPSSLGLP-------------------------PPH
       PtychoderaF --------PITTPYLDHHRDRH--HSALGIS-------------------------PS-
      TrichoplaxAdh YHQQQ--------YRLDNHHHQSLPAFQGIQNTD-------------------------
  BranchiostomaFlo PT-----------------TMPTPALQGIS-----------------------------
    PlatynereisDum N-------------------------LVGL-------------------------QPSA
       MusMuscGsx1 LGRQHS-------VSP------------GV---------------------------AH
       HomoSapGsx1 LGRQHS------AVSP------------GV---------------------------AH
     XenopusTroGsx1 LGRQHS------ASTG-------------I-----------------------------
       DanioRerGsx1 LSRQHS------TSSG-------------V-----------------------------
     XenopusTroGsx2 SQQPS-------------------------------------------------PAM
       DanioRerGsx2 AHQQS-------------------------------------------------PAL
       MusMuscGsx2 SHAHHH------HHPPQHHHHHHQPQQPGSAAAAAAAAAAAAAAA----------AAAL
       HomoSapGsx2 NHAHHH------HHPPQHHHHHHQPQQPGSAAAAAAAA-AAAAA----------AAAL

          CionaInt GCG-SGSSMQQLPAVEHQVSSRVTPDPVT----------DTTSPLCEDDEQGDRKRMRTA
         HydraVulg -SPLHGQ-RAQFSPNY-----VYDLELRHSRQL-QLQH-----QEHETDLYGKSKRIRTA
         ClytiaHem -SPINARARPQFPALY-----VRDLDSRRLQLQQQVQHQRQQQLEEQQGGGSKSKRIRTA
     DrosophilaMel GK-RFKHESSCSPNSS-PLKNHSSGGPVEI-------------TPLINDYADSSKRIRTA
      AnophelesGam -----AGASASTPPTI-P---SASPSPTRSTDLS---------QTYAIDDELSSKRIRTA
      TriboliumCas -----VSVS---PTTP-P---TPSPPPE--------------ERLTPSKDASSKRIXTA
        PatiriaMin S--MSQHTPPSSHTSY-----RPTIDPRRLQYLP---M-ALDHSGMDQDHLSSSKRIRTA
       PtychoderaF ---LPRQATTASPAAF-----RLPVDPRRLQYMP---M-NVDA---NSDNLQSSKRIRTA
      TrichoplaxAdh -------------------------TQQHPDLA---S-DK-TSMKSTSHSSRTKRIRTA
  BranchiostomaFlo YAISRHHTPA---------------TSAFLSLP---F-PSG--VASDGLPRDSRRMRTA
    PlatynereisDum FSPTNSDAP---VSPL------RLADRARLRYLT---V-GGAPVSNAEMEDANGKRIRTA
       MusMuscGsx1 GPAAAAAAAALYQTSY------PLPDPRQFHCIS---V-DSS-SN----QLPSSKRMRTA
       HomoSapGsx1 GPAAAAAAAALYQTSY------PLPDPRQFHCIS---V-DSS-SN----QLPSSKRMRTA
     XenopusTroGsx1 ---NVSHGPALYQAAY------PLPDPRQFHCIS---V-DSS-PS----QLSSSKRMRTA
       DanioRerGsx1 ---NLNSGSGLYQAAY------PVPDPRQFHCIS---I-ENS-SS----QLQSSKRMRTA
     XenopusTroGsx2 VHP--AHSPVCSP-TY------NVTDPRRFHCMT---M-GGAENS----QAQNGKRMRTA
       DanioRerGsx2 AHPGHGHAPVCTPTSF------SVTDPRRYHCLS---L-GASDNS----HIQNGKRMRTA
       MusMuscGsx2 GHP-QHHAPVCAATTY------NMSDPRRFHCLS---M-GGSDTS----QVPNGKRMRTA
       HomoSapGsx2  GHP-QHHAPVCTATTY------NVADPRRFHCLT---M-GGSDAS----QVPNGKRMRTA
                                                                      :*: **
```

```
          CionaInt FTGSQLLELEREFHSDMYLTRLRRIRIAQMLNLSEKQIKIWFQNRRVKKKKGEKPQQDSR
        HydraVulg YTSIQLLELEKEFQNNRYLSRLRRIQIAAILDLTEKQVKIWFQNRRVKWKKDKKG---Y-
         ClytiaHem YTSIQLLELEKEFQNNRYLSRLRRIQIAAMLDLTEKQVKIWFQNRRVKWKKDKKSGFNN-
     DrosophilaMel FTSTQLLELEREFSHNAYLSRLRRIEIANRLRLSEKQVKIWFQNRRVKQKKGGSESPTFN
      AnophelesGam FTSTQLLELEREFAGNMYLTRLRRIEIATRLRLSEKQVKIWFQNRRVKRKKGDAPFGAE-
      TriboliumCas FTSTQLLELEREFASNMYLSRLRRIEIATCLRLSEKQVKIWFQNRRVKYKKEDLPAAAG-
        PatiriaMin FTSTQLLELEREFASNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKYKKETKVQSR--
       PtychoderaF FTSTQLLELEREFAANMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVQYKKRRKTRVR--
     TrichoplaxAdh YTSMQLLELEKEFNSSRYLSRLRRIEIANMLNLSEKQVKIWFQNRRVKWKKDNKLDGQDH
  BranchiostomaFlo FSSTQLLELEREFASNMYLSRLRRIEIATFLNLSEKQVKIWFQNRRVKHKKEARAASGSR
    PlatynereisDum FTSTQLLELEREFSSNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKYKKEGTDSRE--
       MusMuscGsx1 FTSTQLLELEREFASNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEGKGSNHRG
       HomoSapGsx1 FTSTQLLELEREFASNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEGKGSNHRG
    XenopusTroGsx1 FTSTQLLELEREFASNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEGKSSTHRA
      DanioRerGsx1 FTSTQLLELEREFTSNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEGKSGSHRT
    XenopusTroGsx2 FSSTQLLELEREFSSNMYLSRLRRIEIATYLSLSEKQVKIWFQNRRVKHKKENKGSQRNT
      DanioRerGsx2 FTSTQLLELEREFSSNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEGKGTQRSA
       MusMuscGsx2 FTSTQLLELEREFSSNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEGKGASRNN
       HomoSapGsx2 FTSTQLLELEREFSSNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEGKGTQRNS
                   ::.  ******:** . **:*****.**  * *:*** :*********:**
```

Figure 3.1 Clustal O aligment of protein sequence of Gsx family members. Gsx sequences from representatives of the main animal phyla have been considered. Blue box: eh1-like domain. Green box: PstI domain, which is insect specific. Yellow box: *Drosophila* activation domain. Pink box: homeodomain.
Red residues within the eh1 domain indicate the invariant residues in the consensus motif; purple amino acids indicate residues whose biochemical properties are generally conserved in the Eh1 motives so far identified. Green residues indicate amino acids highly conserved in the Eh1 proteins domain of Gsx proteins, which were not described as general features of the other Eh1-containing transcription factors previously analyzed.

Ci-GSX EXPRESSION PATTERN

3.2 Refining *Ci-gsx* expression pattern

Previous studies (Hudson and Lemaire, 2001) and data available on ANISEED database reported *Ci-gsx* expression starting from gastrula stage in the couple of cells a9.33, in row III of the neural plate (***figure 1.12***). My studies were firstly devoted to further confirm these data, then to better analyze *Ci-gsx* expression at different developmental times and define the territories in which the transcript is expressed. To this aim, I performed single and double *in situ* hybridization experiments, comparing *Ci-gsx* expression with different markers of *C. intestinalis* CNS lineages, at different embryonic stages.

3.2.1 Confirming previous data: Ci-gsx expression pattern

Whole Mount *in situ* Hybridization experiments (WMISH) to localize *Ci-gsx* transcript were carried out in embryos at different developmental stages. Results obtained from gastrula to late

tailbud are shown in ***figure 3.2***. The expression of *Ci-gsx* starts at gastrula stage, in two cells, likely the a9.33 couple, as previously observed (Hudson and Lemaire, 2001; ***figure 3.2 A*** compared to ***B***). At neurula stage *Ci-gsx* appears also in the two blastomeres between the a9.33 pair, likely the a9.37 cells (***figure 3.2 C***). Later on in development, the signal becomes wider and distributed in two symmetrical stripes along the AP axis, in the anterior part of the embryo (***figure 3.2 D*** and ***E***), up to late tailbud stage (***figure 3.2 F***). After this stage, the signal remains in the developing sensory vesicle, but becomes weaker and weaker, hardly distinguishable from the background.

This preliminary result indicated that both the proposed precursors of photoreceptor cells are likely to be be marked by *Ci-gsx* at the time where their fate starts to be restricted.

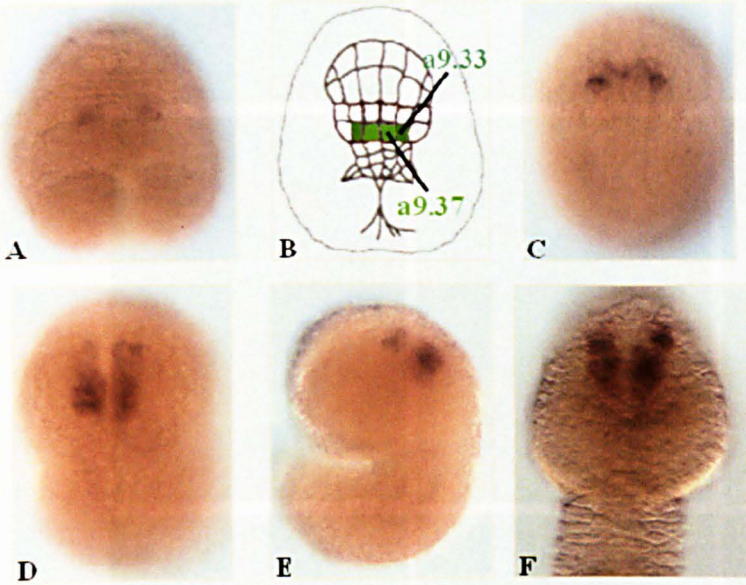


Figure 3.2 *Ci-gsx* expression pattern during *Ciona intestinalis* development. WMISH with *Ci-gsx* probe on embryos at different developmental stages. **A.** Gastrula, dorsal view. **B.** Schematic representation of the gastrula stage; a9.33 cell pair is depicted in dark green and a9.37 couple in light green. **C.** Neurula, dorsal view. **D.** Middle tailbud, dorsal view. **E.** Middle tailbud, lateral view. **F.** Late tailbud, anterior region, dorsal view.

3.2.2 Comparison of Gi-gsx and Ci-Six 3/6 expression pattern

Ci-Six3/6 is a pro-ortholog of vertebrates *Six3* and *Six6* (Wada *et al.*, 2003), specifically expressed, *in Ciona*, in the row IV of the neural plate at late gastrula stage (Moret *et al.*, 2005b). The a9.33 and a9.37 cell couples belong to row III, immediately below the row IV (***figure 3.3 A, B***). Therefore, I used *Ci-Six3/6* marker together with *Ci-gsx* in double *in situ* experiments, in order to characterize more precisely the *Ci-gsx* positive cells at the earliest stages of expression.

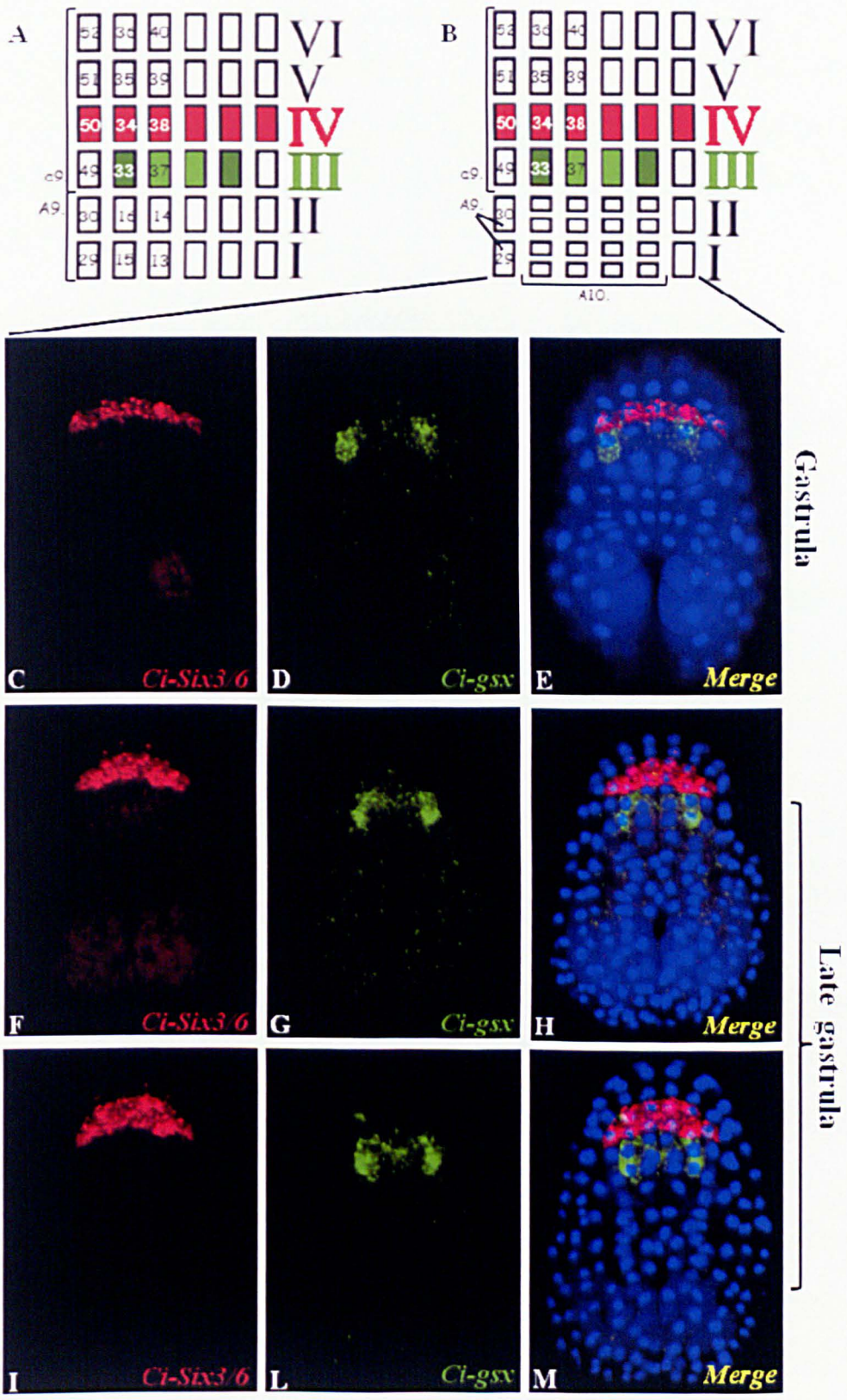


Figure 3.3 Positional relationship of *Ci-gsx* and *Ci-Six3/6* gene expression. A, B: Schemes of the neural plate at two different moment of the gastrula stage. The fourth row, marked by *Ci-Six3/6* is indicated in red; a9.33 and a9.37 cell couples in the third row are indicated in dark and light green respectively. **C-M:** Double WMISH at developmental stages indicated on the right of the picture. **C, F, I**: staining for *Ci-Six3/6* probe. **D, G, L**: staining for *Ci-gsx* probe. **E, H, M**: merge of the respective previous single images, with the addition of the DAPI nuclear staining (in blue). All the embryos are in dorsal view.

The analysis clearly confirmed that *Ci-gsx* is firstly expressed in a9.33 cells at gastrula stage (***figure 3.3 C-E***), while the signal in the a9.37 lineage appears later, at the late gastrula stage, just after their division, and seems to be localized only into the posterior progeny of a9.37 couple, whereas the anterior progeny, just below *Ci-Six3/6* territory, results unstained (***figure 3.3 F-M***).

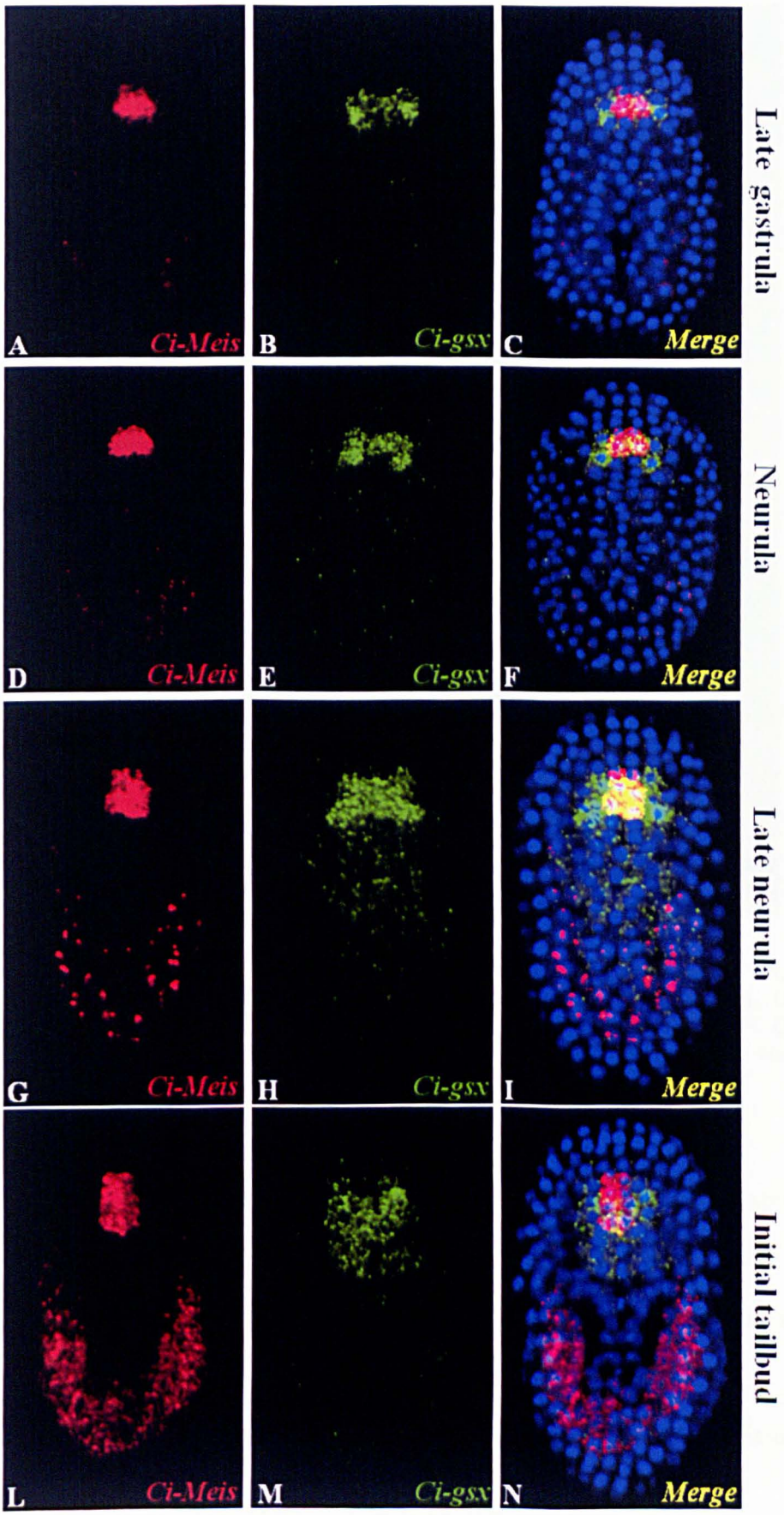
3.2.3 Comparison of Gi-gsx, Ci-Meis and Ci-Ptf1a expression pattern

Ci-Meis and *Ci-Ptf1a* are considered two markers of the dopamine (DA)-synthetizing cells which are part of the coronet cells of the brain vesicle (ANISEED data; Moret *et al.*, 2005a; Razy-Krajka *et al.*, 2012). *Ci-Meis* and *Ci-Ptf1a* were thus used, each one in combination with *Ci-gsx*, to outline any overlapping expression pattern during *Ciona* development. Here I present the data concerning *Ci-gsx/Ci-Meis*, since, based on my data, the expression profile of *Ci-Meis* and *Ci-Ptf1a* were almost identical (from neurula stage *Ci-Meis* expression is wider, being expanded anteriorly with respect to *Ci-Ptf1a*; data not shown).

At the late gastrula stage *Ci-Meis* transcript demarcates the whole a9.37 progeny (a10.73 and a10.74 pairs; ***figure 3.4 A-C, D-F***) and *Ci-gsx* co-localize with *Ci-Meis* in the posterior pair (***figure 3.4 G-I***).

Ci-gsx co-localize with *Ci-Meis*, in its medial posterior domain, also in the subsequent embryonic stages (***figure 3.4 G-I***), up to initial tailbud stage. Then, *Ci-gsx* and *Ci-Meis* expression territories become slowly not anymore superimposable (***figure 3.4 L-N***) up to appear mutually exclusive (***figure 3.4 O-Q, R-T***).

During neural tube closure, in fact, *Ci-Meis* signal persists in the medial region and becomes slightly internalized. Conversely, *Ci-gsx* disappears from the medial area and becomes evident in two separate and bilateral domains localized anterior and posterior to the *Ci-Meis* median territory (***figure 3.4 O-Q, R-T***). These findings suggest that the expression of *Ci-gsx* is present also in part of the a9.37 progeny and is kept in this territory up to the early tailbud stage.


A
Ci-Meis
B
Ci-gsx
C
Merge
Late gastrula
D
Ci-Meis
E
Ci-gsx
F
Merge
Neurula
G
Ci-Meis
H
Ci-gsx
I
Merge
Late neurula
L
Ci-Meis
M
Ci-gsx
N
Merge
Initial tailbud

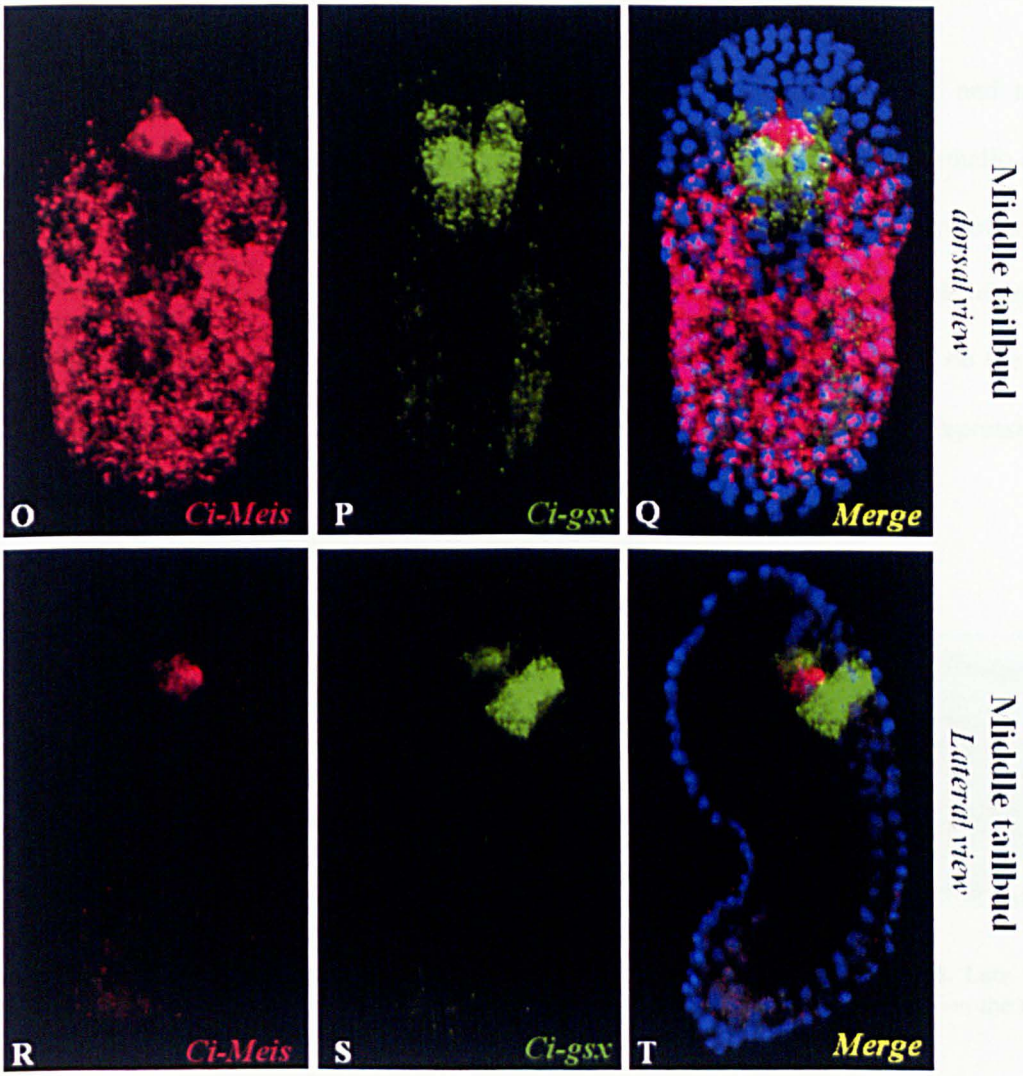


Figure 3.4 Positional relationship of *Ci-gsx* and *Ci-Meis* gene expression. Double WMISH at different embryonic stages, reported next to each picture row. **A**, **D**, **G**, **L**, **O**, **R**: staining for *Ci-Meis* probe. **B**, **E**, **H**, **M**, **P**, **S**: staining for *Ci-gsx* probe. **C**, **F**, **I**, **N**, **Q**, **T**: merge of the respective previous single images, with the addition of the DAPI nuclear staining (in blue). A-Q: dorsal view. R-T: lateral view.

It is possible to conclude that at least in early embryonic stages, *Ci-gsx* is expressed in the subpopulation of the DA-synthetizing cells precursors, that will become coronet cells in the ventral part of the sensory vesicle. This result is highly interesting in an evolutionary perspective, as discussed in the next chapter.

It is also intriguing to note that the posterior expansion of *Ci-gsx* domain from the initial tailbud stage is likely to correspond to the capital A cell lineage.

3.2.4 Comparison of Ci-gsx and Ci-Rx expression pattern

The homeobox transcription factor *Ci-Rx* plays a key role in the development and in the differentiation of the photoreceptor cells and pigment cell of the ocellus organ (D'Aniello *et al.*, 2006: ***figure 3.5***). *Ci-Rx* mRNA first appears in the early tailbud stage in two groups of cells in the most anterior region of the developing sensory vesicle (***figure 3.5 A***). As development proceeds the left signal disappears while the signal persists and becomes stronger on the right side, where photoreceptors form (***figure 3.5 B, C***). At the larval stage *Ci-Rx* expression is visible around ocellus, just in the photoreceptor cells territories (***figure 3.5 D***).

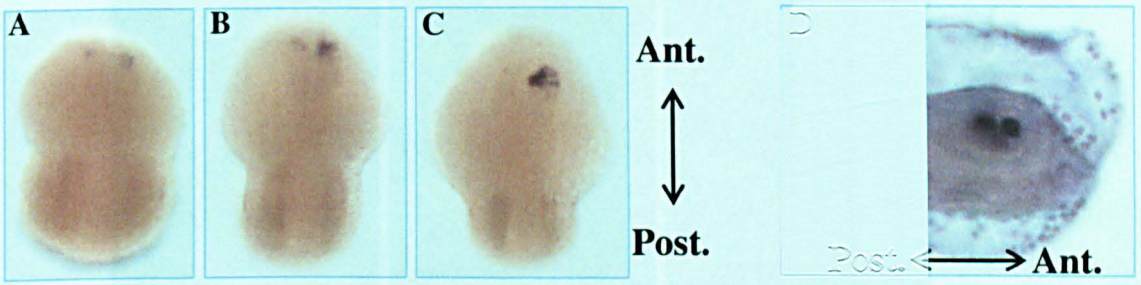


Figure 3.5 *Ci-Rx* expression pattern (D'Aniello *et al.*, 2006). **A.** Early tailbud stage**B.** Late tailbud stage I. **C.** Late tailbud stage II. **D.** Larva in lateral view. A-C: Dorsal view. Anterior in on the top; D: anterior is on the right.

Given its peculiarity as photoreceptor marker, *Ci-Rx* has thus been used in double *in situ* hybridization experiments to define the expression territories of *Ci-gsx* from the early tailbud stage. The data clearly demonstrated that *Ci-gsx* and *Ci-Rx* perfectly co-localize in the most anterior part of *Ci-gsx* expression domain, since the early tailbud stage and up to the late tailbud stage (***figure 3.6***), thus supporting a potential role of *Ci-gsx* in photoreceptor cells formation.

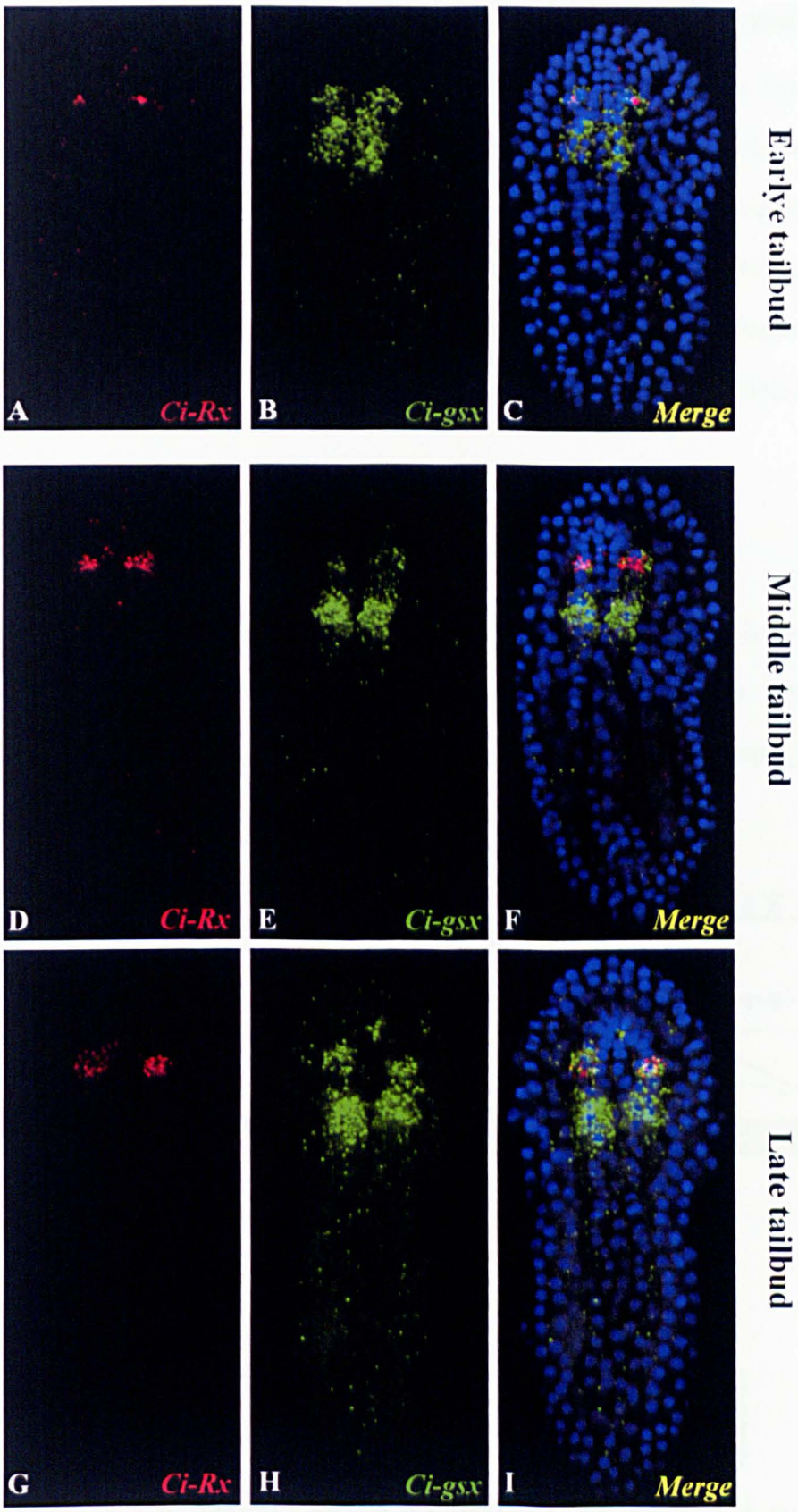


Figure 3.6 Positional relationship of *Ci-gsx* and *Ci-Rx* gene expression. Double WMISH at different embryonic stages, reported next to each picture row. All the embryos are in dorsal view. **A**, **D**, **G**: staining for *Ci-Rx* probe. **B**, **E**, **H**: staining for *Ci-gsx* probe. **C**, **F**, **I**: merge of the respective previous single images, with the addition of the DAPI nuclear staining (in blue).

DEFINING Ci-GSX ROLE IN PHOTORECEPTOR CELLS PRECURSORS

Based on the data so far illustrated, I decided to continue my work mainly following two directions: 1) investigate the role of Ci-gsx in photoreceptor cells differentiation; 2) gain insights into the regulatory mechanisms underlying *Ci-gsx* specific expression pattern. It was preparatory for both approaches the isolation of *Ci-gsx* regulatory region. The *Ci-gsx* promoter has been indeed instrumental for lineage specific targeted perturbation of endogenous *Ci-gsx* and, in parallel, for the identification of putative *cis*-acting elements responsible for *Ci-gsx* regulation.

3.3 Isolation of Ci-gsx regulatory region

To isolate the sequence able to activate *Ci-gsx* tissue specific expression, a 2.8 kb genomic fragment was amplified by PCR from *Ciona* genomic DNA, using the most suitable oligonucleotides designed on the sequence available on JGI1 (http://genome.jgi-psf.org/ciona4/ciona4.home.html; ***figure 3.7 A, B***).

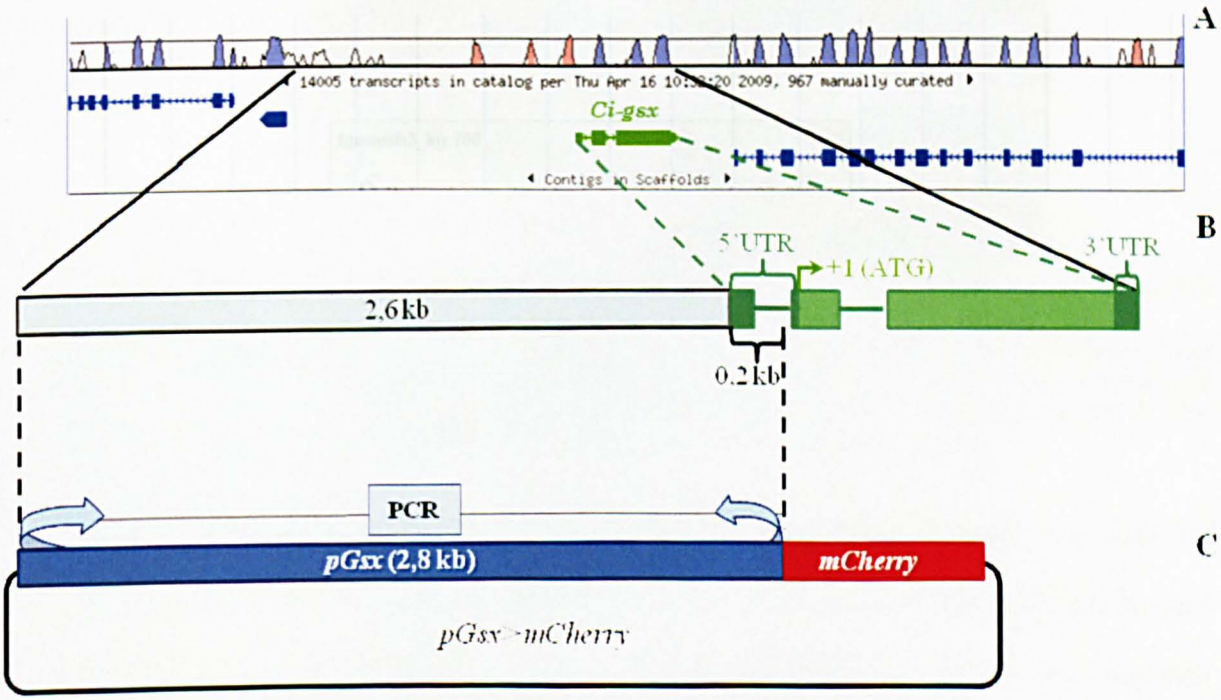


Figure 3.7 Isolation of a *cis*-regulatory element required for *Ci-gsx* activation in the endogenous territories. A. Visual representation of the genomic region containing *Ci-gsx* coding sequence (in green). Adapted from JGI database. **B.** Detailed representation of the genomic region containing *Ci-gsx* gene (in green) and the upstream putative regulatory fragment (light blue bar). *Ci-gsx* CDS is represented with light green bars; 5'UTR and 3'UTR are represented by dark green bars; dark green line represents introns. **C.** Schematic representation of *pGsx>mCherry* construct, in which *pGsx* fragment, previously amplified by PCR (blue bar), is cloned upstream from the reporter gene mCherry coding sequence (red bar). As indicated in the picture, *pGsx* also contains part of the 5'UTR and of an intron.

The amplified region, named *pGsx*, extends from the position -2817 to -35 upstream of Ci-gsx putative translation start site; it also contains a part of the 5'UTR of *Ci-gsx* transcript, (including an intron), from the position -264 to -35 upstream from the ATG starting codon of the gene (***figure 3.7 B, C***).

This putative regulatory region was cloned upstream from *mCherry* reporter gene, to create the construct *pGsx>mCherry* (***figure 3.7 C***) that was introduced into *C. intestinalis* fertilized eggs *via* electroporation; the embryos were then checked, under fluorescent microscopy, during their development (***figure 3.8***).

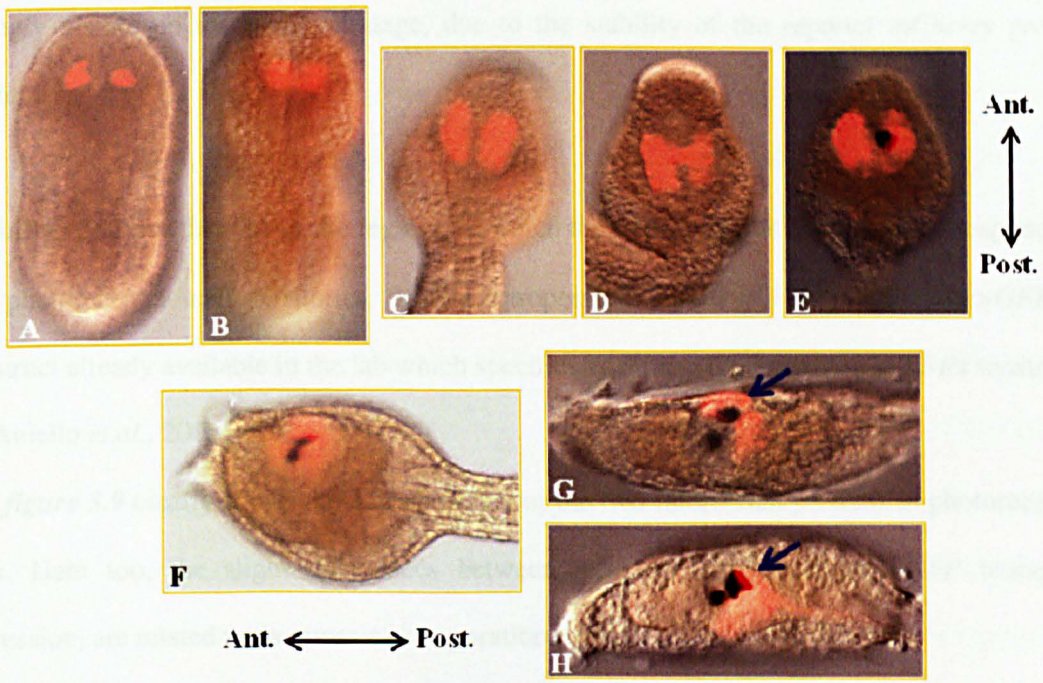


Figure 3.8 In vivo analysis of *pGsx>mCherry* construct. Merged bright-field/fluorescent images are reported. In pictures **A-E** dorsal view, anterior is on the top; **F-H** lateral view, anterior is on the left.
A. Late neurula stage, **B.** Middle tailbud stage, **C,D.** Late tailbud stages, anterior part of the embryo. Note the different expression of the reporter, due to the mosaic incorporation of the transgene. **E.** Late tailbud II stage, anterior part of the embryo. **F.** Late tailbud III stage, anterior part of the embryo, dorso-lateral view. **G, H.** Trunk of transgenic larvae; blue arrows indicate the photoreceptor cells region.

Almost 90% of the electroporated embryos showed a strong transgene expression, from late neurula (***figure 3.8 A***) to late tailbud stage, specifically in two symmetrical domains quite comparable with that of endogenous *Ci-gsx* (***figure 3.8 B-E***). The fluorescence was kept bilateral up to the larval stage, with a strong signal clearly detectable in the photoreceptor cells region and/or in the ventral part of the sensory vesicle (***figure 3.8 F-H;*** blue arrows indicate the photoreceptor cells).

The variations in the territories expressing mCherry, among different embryos (***figure 3.8 C*** compared with ***D***, and ***G*** compared with ***H***), are in agreement with the mosaic incorporation of transgenes in ascidians, in the sense that the exogenous plasmid can possibly not be equally inherited in all the blastomeres of the lineage during development.

Concerning the observed temporal delay in the appearance of the fluorescent signal (late neurula) with respect to the expression of the endogenous *Ci-gsx* (gastrula), this phenomenon was expected, since it is related to the time necessary to accumulate a detectable amount of fluorescent protein. Furthermore, differently from *Ci-gsx* transcript, the fluorescence was strongly visible up to the larval stage, due to the stability of the reporter *mCherry* protein product.

To further confirm that one of the regions, in which the transgene was expressed, corresponds to the photoreceptor cell territories, I co-electroporated *pGsx>mCherry* with *pRx>GFP*, a construct already available in the lab which specifically directs GFP reporter in *Ci-Rx* territories (D'Aniello *et al.*, 2006).

The ***figure 3.9*** clearly shows the co-expression of the two fluorescent proteins in photoreceptor cells. Here too, the slight differences between *pGsx>mCherry* and *pRx>GFP* transgene expression, are related to the mosaic incorporation.

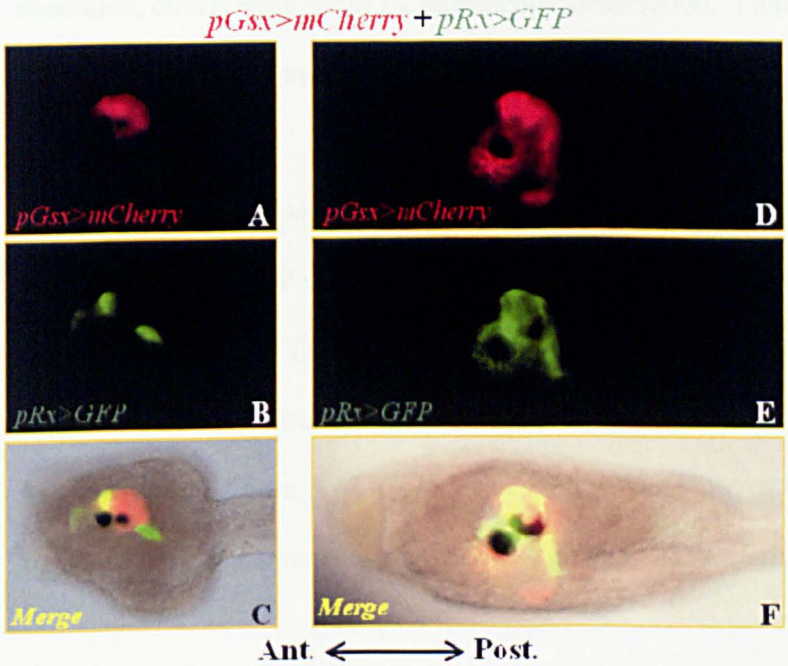


Figure 3.9 Co-electroporation experiment of *pGsx>mCherry* and *pRx>GFP*. Anterior is on the left. **A-C.** Late tailbud III embryos, anterior region. D-F. Larvae, trunk region. **A, D**. Red fluorescence driven by *pGsx* promoter. **B, E**. Green fluorescence driven by *pRx* promoter. **C, F**. Merge of the two previous respective channels, plus bright-field. Note the yellow regions, indicating co-localization of the two reporters.

It was therefore possible to conclude that the tested *pGsx* region contained all the *cis*-regulatory elements necessary for the proper spatio-temporal expression of the downstream gene, thus representing a useful tool in order to carry out promoter-guided lineage-specific interfering strategies.

3.4 Perturbing the activity of the endogenous Ci-gsx

To test the role of Ci-gsx in the formation of photoreceptor cells, I prepared two chimeric constructs potentially able to interfere with Ci-gsx activity. *GsxHD>VP16*, a constitutively active form of Ci-gsx, in which the region coding for Ci-gsx homeodomain was fused in frame with VP16 trans-activation domain, and *GsxHD>WRPW*, a constitutively repressor form of Ci-gsx, in which the region coding for Ci-gsx homeodomain was fused in frame with WRPW repressor motif (Davidson *et al.*, 2006). These interfering forms of *Ci-gsx* (*GsxHD>VP16* and *GsxHD>WRPW)* were selectively expressed in *Ci-gsx* territories using the *pGsx* promoter previously identified (constructs *pGsx>GsxHD>VP16* and *pGsx>GsxHD>WRPW).* The rationale behind this experiment was that one of the two chimeric produced proteins should contrast the activity of the endogenous Ci-gsx, that could in turn induce an altered developmental program in the transgenic embryos, hopefully at the level of photoreceptor cells. However, putative defects in photoreceptor cells formation do not results into a morphological aberration, clearly identifiable by microscopic observation. I thus used three different strategies to study the "hidden" phenotypes:

As a first approach, embryos electroporated with *pGsx>GsxHD>VP16* and *pGsx>GsxHD>WRPW* were tested by *in situ* hybridization experiments in order to check the possible alteration of *Ci-Rx* expression, in comparison with control embryos, electroporated with *pGsx>LacZ* construct.

As shown in the ***figure 3.10,*** *Ci-Rx* expression was not altered in embryos electroporated with the constitutive repressor construct, *pGsx>GsxHD>WRPW* (***figure 3.10 B*** compared to ***A***).

Conversely, *Ci-Rx* signal was strongly reduced, often completely absent, in almost all the embryos electroporated with the constitutive active construct, *pGsx>GsxHD>VP16* (***figure 3.10 C***, compared to ***A***).

The embryos were observed at the middle tailbud stage, when *Ci-Rx* expression is robust.

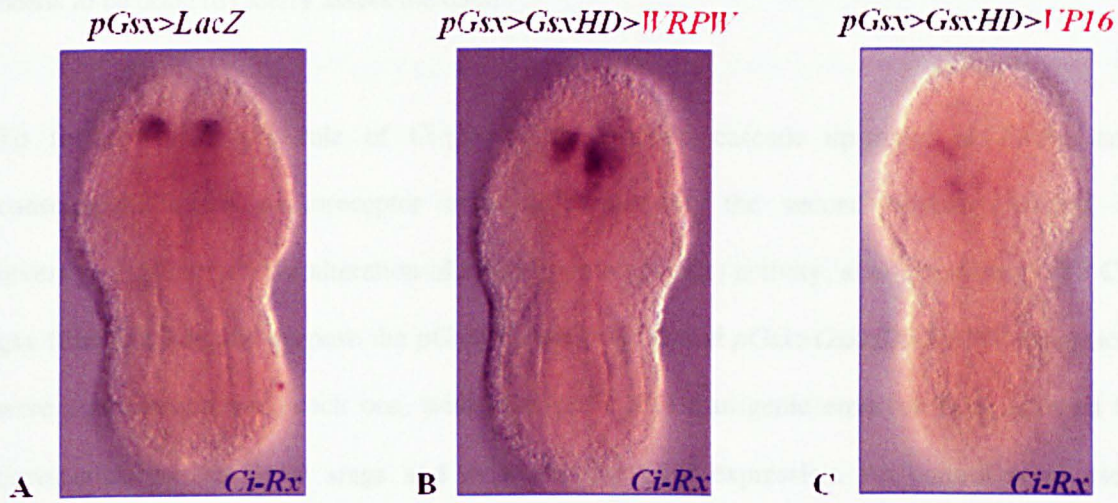


Figure 3.10 WMISH hybridization with *Ci-Rx* probe on transgenic embryos electroporated with Ci-gsx interfering constructs. Middle tailbud embryos, electroporated with (**A**) control *pGsx>LacZ*; (**B**) *pGsx>GsxHD>WRPW*; (**C**) *pGsx>GsxHD>VP16* constructs. All the embryos are in dorsal view, anterior in on the top.
The experiment has been performed in three different batches of animals, and for each condition at least 25 embryos were analyzed.

The observation of embryos at later embryonic stages revealed an even more intriguing phenotype. Transgenic *GsxHD>VP16* late tailbud embryos, as their younger relatives middle tailbud embryos, did not show *Ci-Rx* expression (***figure 3.11 C***, compared to ***A***).

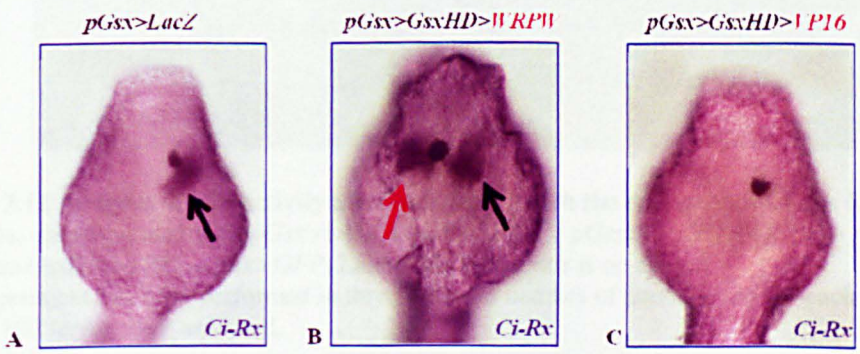


Figure 3.11 WMISH hybridization with *Ci-Rx* probe on transgenic embryos electroporated with Ci-gsx interfering constructs. Late tailbud II embryos, electroporated with (**A**) control *pGsx>LacZ*; (**B**) *pGsx>GsxHD>WRPW*; (**C**) *pGsx>GsxHD>VP16* constructs. Black arrows indicate the wild type *Ci-Rx* expression. The red arrow in picture B highlights the expansion of *Ci-Rx* expression in this experimental condition. All the embryos are in dorsal view, anterior in on the top. The experiment has been performed in three different batches of animals, and for each condition at least 25 embryos were analyzed.

Conversely, in some *GsxHD>WRPW* late tailbud embryos (around 40% of embryos) *Ci-Rx* expression was kept bilateral, compared to the control embryos in which *Ci-Rx* mRNA was restricted to the right side of the developing sensory vesicle (***figure 3.11 B*** red arrow, compared to ***A***), as it normally does. However these results are very preliminary and more experiments needs to be done to clearly assess the data.

To further verify the role of Ci-gsx in the genetic cascade upstream of *Ci-Rx*, and consequentially on photoreceptor cells differentiation, the second approach aimed at investigating the putative alteration of *Ci-Rx* promoter (*pRx*) activity, after interfering with Ci-gsx function. For this purpose the *pGsx>GsxHD>VP16* and *pGsx>GsxHD>WRPW* constructs were co-electroporated, each one, with *pRx>GFP*. The transgenic embryos were allowed to develop up to the larval stage and evaluated for GFP expression. As control were used transgenic larvae co-electroporated with *pGsx>LacZ* + *pRx>GFP*, which showed GFP expression in 50% of specimens (***figure 3.12 A***).

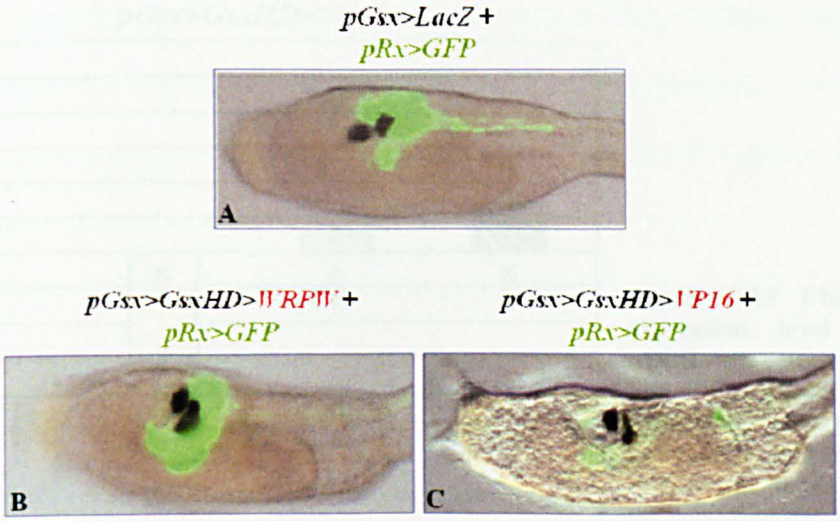


Figure 3.12 Analysis of *pRx* activity after interfering with the endogenous Ci-gsx function.
A. Larva, electroporated with *pGsx>LacZ* + *pRx>GFP*, **B.** *pGsx>GsxHD>WRPW* + *pRx>GFP*
C. *pGsx>GsxHD>VP16* + *pRx>GFP*. Lateral view, anterior is on the left.
The experiment has been performed in three different batches of animals, and for each condition at least 100 larvae were analyzed.

The percentage of larvae expressing GFP dropped down to 15%, respect to the controls, after the electroporation of *pRx>GFP* plus *pGsx>GsxHD>VP16* (***figure 3.12 C***). On the contrary, fluorescence was strongly detectable in 90% of the larvae in which *pRx>GFP* was co-

electroporated with *pGsx>GsxHD>WRPW* (***figure 3.12 B***). In accordance with my previous observations on *Ci-Rx* expression.

In the third approach, the possible changes in the expression levels of *Ci-Rx* transcripts, in embryos electroporated with the two chimeric forms of Ci-gsx, were evaluated by Real Time-qPCR. As control were used transgenic embryos electroporated with *pGsx>mCherry*. All the qPCR experiments have been done in triplicate, on three different batches of embryos. Total RNA was extracted from whole embryos at late tailbud stage, when endogenous *Ci-Rx* expression is robust. The data have been normalized using *Ci-Rps27* and *Ci-Bra* as reference genes and a fold change greater than or equal to 1,5 (from -1,5 to +1,5 with respect to the control population) has been considered significant. The results reported in ***figure 3.13*** indicated that the trend of *Ci-Rx* fluctuations parallels the previous results, with a downregulation of the gene in embryos electroporated with *pGsx>GsxHD>VP16* and an upregulation in embryos electroporated with *pGsx>GsxHD>WRPW*.

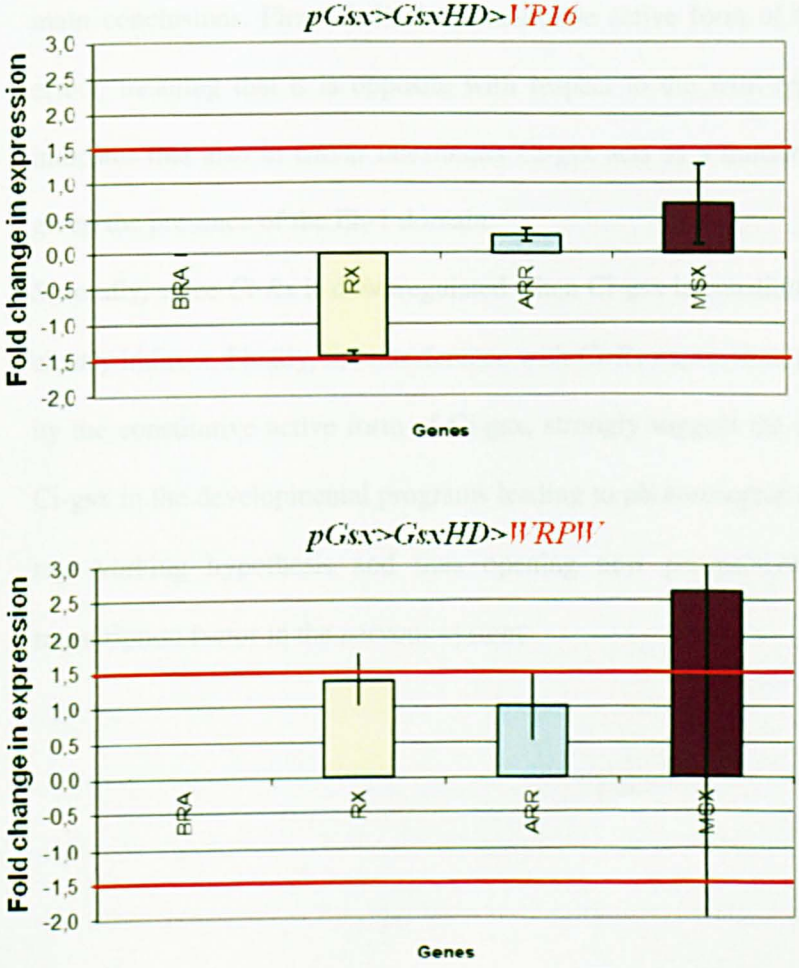


Figure 3.13 Change in gene expression level assessed by qPCR on *pGsx>GsxHD>VP16* and *pGsx>GsxHD>WRPW* transgenic embryos. Data are expressed as fold change difference in transcript levels with respect to the control larvae. Fold differences greater than or equal to ± 1,5 (indicated by red lines) are considered significant.

In order to have an additional proof of photoreceptor cells alterations, I also attempted to analyze *Ci-Arr* variations in the same conditions, being *Ci-Arr* a marker for differentiated photoreceptor cells (Nakagawa *et al.*, 2002). As shown in the graph, the change is not significant in embryos expressing the constitutive active Ci-gsx, and is barely relevant in embryos electroporated with the constitutive repressor form (***figure 3.13***, light blue bars).

Finally, given that *Msx* is a known Gsx target in *Drosophila* nervous system, I included the analysis of *Ciona* ortholog gene in my study. Unfortunately, the analysis of *Ci-msxb* expression was not informative, because of the low fold change of the transcript in one case (*GsxHD>VP16*), and the variability of the results in the replicates in the other one (*GsxHD>WRPW*; ***figure 3.13***, purple bars).

3.4.1 Preliminary conclusions (Part I)

The targeted perturbation of the endogenous Ci-gsx function, obtained by expressing a constitutively active and repressive form of the transcription factor, allowed me to draw three main conclusions. First of all, the constitutive active form of Ci-gsx exhibits an antimorphic effect, meaning that it is opposite with respect to the wild type condition. This observation indicates that also in *Ciona intestinalis* Ci-gsx acts as a transcriptional repressor, as supposed given the presence of the Eh-1 domain.

Secondly, since *Ci-Rx* is downregulated when Ci-gsx is constitutively active, their interaction is clearly indirect. Finally, the interference with *Ci-Rx* expression, and thus Ci-Rx activity, exerted by the constitutive active form of Ci-gsx, strongly suggest the involvement of the endogenous Ci-gsx in the developmental programs leading to photoreceptor cells differentiation, confirming my working hypothesis and thus opening new perspectives about the function of this transcription factor in the nervous system.

STUDIES ON Ci-GSX REGULATORY REGION

As already mentioned, the second main goal of my PhD work has been the study of *Ci-gsx* transcriptional regulation, through the identification of the minimum sequence responsible for the tissue-specific expression of the gene.

3.5 Dissection of *Ci-gsx* promoter

Starting from the previously isolated 2.8kb *pGsx* fragment, which was able to drive a strong expression of the reporter almost in 90% of the tested embryos, I narrowed down this regulatory region by using bioinformatics tools. In particular, the version 2 of *C. intestinalis* genomic sequence, JGI2, (http://genome.jgi-psf.org/Cioin2/Cioin2.info.html) automatically performs a genetic comparison of all annotated genes, as well as intergenic regions, between *C. intestinalis* and the closely related species *C. savignyi*. This comparison is obtained through the mVISTA genomic tool (Mayor *et al.*, 2000) and is very useful to find conserved regulatory sequences functionally relevant.

Figure 3.14 (top left) depicts the genomic comparison of the *Ci-gsx* sequence of *C. intestinalis* and *C. savignyi*. Conserved non-coding regions are represented by pink peaks in the plot, while conserved coding regions are represented by blue peaks.

Three conserved regions between the two species are present in *pGsx* promoter, and I named them I, II and III, from the more distal to the more proximal to *Ci-gsx* coding sequence. Based on this, my next strategy was the isolation by PCR of the genomic region laying immediately upstream from *Ci-gsx* gene and containing the three peaks of homology with *C. savigny*, thus excluding the non-conserved part. This fragment of 1.9 kb (*pGsxA*), which extends from the position -1988 to -35 from the translation start site, showed an activity similar to that of *pGsx* (***figure 3.14***). The removal of the 5'UTR (construct *pGsxB*, 1.7 kb, which extends from the position -1988 to -264) from Ci-gsx starting codon, did not result in any decrease of activity (***figure 3.14***, green box). Thus the genomic region -1988 to -264, that includes the three peaks of similarity with *C. savignyi*, contained all the *cis*-elements necessary for the robust and complete expression of the reporter gene (***figure 3.14***, on the right).

Many other constructs have then been prepared, in the attempt to understand the biological significance of the three conserved regions. As indicated in the scheme presented in ***figure 3.14***, different combinations of conserved/non-conserved (with *C. savignyi*) sequences have been tested *in vivo*.

The constructs *pGsxI* and *pGsxD*, which lacks the region III closest to *Ci-gsx* transcription start site, were cloned into a plasmid containing the Epstein Barr virus TATA (E1bTATA) basal promoter (Leong *et al.*, 1988; Parks *et al.*, 1988) upstream from mCherry. This short sequence is transcriptionally inactive *per se*, but works finely in *Ciona* as previously tested in the lab (Squarzoni *et al.*, 2011; ***figure 3.14***, purple asterisks).

All the plasmids have been tested in transgenesis *via* electroporation experiments and the embryos were allowed to develop up lo larval stage to be checked at the microscope.

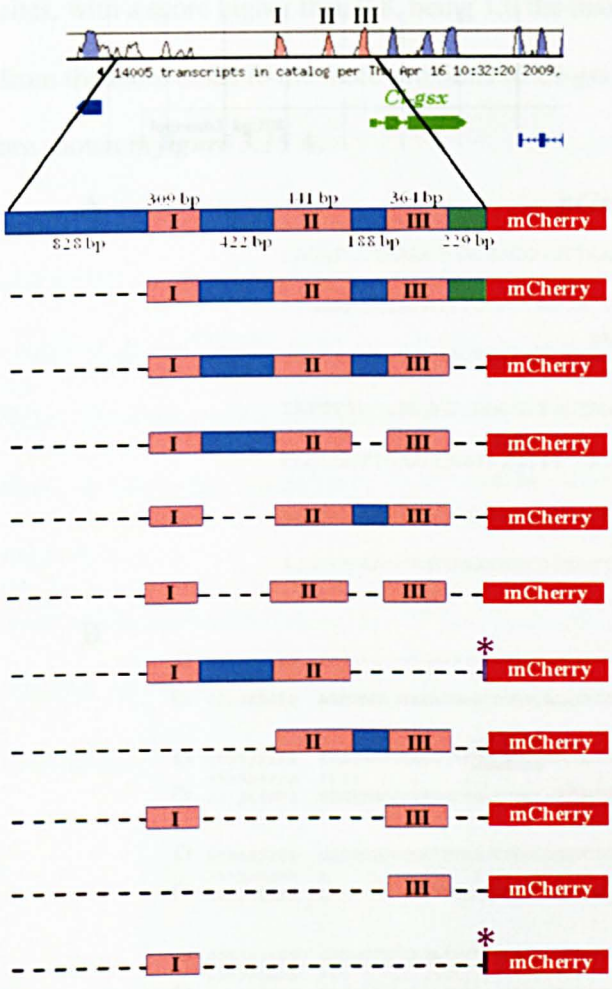
A detailed analysis of territories and percentage of expression of the reporter has been done. All the tested constructs, excepted *pGsxI>mCherry* and *pGsxD>mCherry*, were transcriptionally active and mCherry signal, when present, was always confined to a territory corresponding to *Ci-gsx* expression domain; however, little variations in the percentage of fluorescent larvae were detected as reported in the scheme in ***figure 3.14*** (at the right).

Observing the fluctuations of percentages from one case to another, my main conclusion is that *Ci-gsx* transcriptional regulation is likely to be complex, being not related to a single small transcriptional module. Probably region I, II and III, as well as the non-conserved ones, are all necessary to allow the robust expression that characterize this strong promoter, since by eliminating each one of them, a reduction in the percentage of fluorescent larvae has been found.

Actually, even if I did not precisely quantify this observation, the main difference lied in the "brightness" of the fluorescent signal, which was always fainter with respect to the one I usually observed in *pGsx>mCherry* electroporated larvae, thus confirming that many *cis*-acting elements are necessary to enhance the expression of the reporter gene.

However, the "core" of *Ci-gsx* promoter seemed located in region III, since its absence resulted in loss of promoter activity, as observed with *pGsxI>mCherry* and *pGsxD>mCherry* constructs.

The fragment *pGsxIII* (364 bp), which actually corresponds to this third conserved region, extends from -629 to -283 upstream from the translation starting codon, and it was able to drive the expression of the reporter mCherry in 70% of the larvae (column "+++" added to "+" in ***figure 3.14***). For the next studies, I focused my attention on this fragment that seems to contain all the information necessary to recapitulate the expression of the endogenous gene.



n ≥ 100

PLASMID NAME	+++	+	No
pGsx (2,8 kb)	90.0%	0.0%	10.0%
pGsxA (1,95 kb)	90.0%	0.0%	10.0%
pGsxB (1,7 kb)	90.0%	0.0%	10.0%
pGsxE (1,5 kb)	60.0%	20.0%	20.0%
pGsxF (1,3 kb)	60.0%	20.0%	20.0%
pGsxU (1,1 kb)	70.0%	15.0%	15.0%
pGsxD (1,2 kb)	0.0%	0.0%	**100.0%**
pGsxC (993 bp)	50.0%	35.0%	15.0%
pGsxI+III (673 bp)	45.0%	30.0%	25.0%
pGsxIII (364 bp)	40.0%	30.0%	30.0%
pGsxI (309 bp)	0.0%	0.0%	**100.0%**

Figure 3.14 Summarizing scheme of the results obtained from the *Ci-gsx* promoter region analysis. *On the top **left***: visual representation of the genomic region containing *Ci-gsx* coding sequence (in green). Adapted from JGI database. Immediately below the different tested fragment are represented. The numbers in correspondence of *pGsx* indicate the length of each conserved or non-conserved fragment.
In all the cases pink boxes indicate the conserved regions between *C. intestinalis and C. savignyi* and the blue boxes indicate the non conserved region; the green boxes in pGsx and *pGsxA* regions indicate the part of 5'UTR of the endogenous *Ci-gsx*. The purple asterisks in *pGsxD* and *pGsxI* fragment indicate the presence of the minimal E1bTATA region, inserted by PCR. The red boxes indicate the coding sequence of the reporter gene *mCherry*.
The table ***on the right*** reports the percentage of expression for each tested fragment, whose name and length is indicated in the first column ("Plasmid"). Column "+++" : percentages of larvae in which the expression of the reporter is full and robust with respect to the whole *pGsx* region. Column "+": percentages of larvae in which the fluorescence is detectable, but the signal seems less bright. Column "No": percentage of larvae in which the fluorescence is not detectable. Note that pGsxD and pGsxI regions are not transcriptionally active. The fragment *pGsxIII*, whose row is highlighted by the purple bar, is the smallest region able to drive the expression of the downstream gene.

3.6 Searching for putative transcription factor binding sites in pGsxIII regulatory region (part I)

In order to identify putative *trans*-acting factors that can efficiently regulate *Ci-gsx* expression, I performed an *in silico* analysis on *pGsxIII* fragment using JASPAR database (http://jaspar.genereg.net/). The software revealed the presence of three consensus Sox binding sites, with a score higher than 0.8, being 1.0 the maximum score. I named them sx0, sx1 and sx2 from the more distal to the more proximal to *Ci-gsx* gene. Their positions and the relative scores are shown in ***figure 3.15 A.***

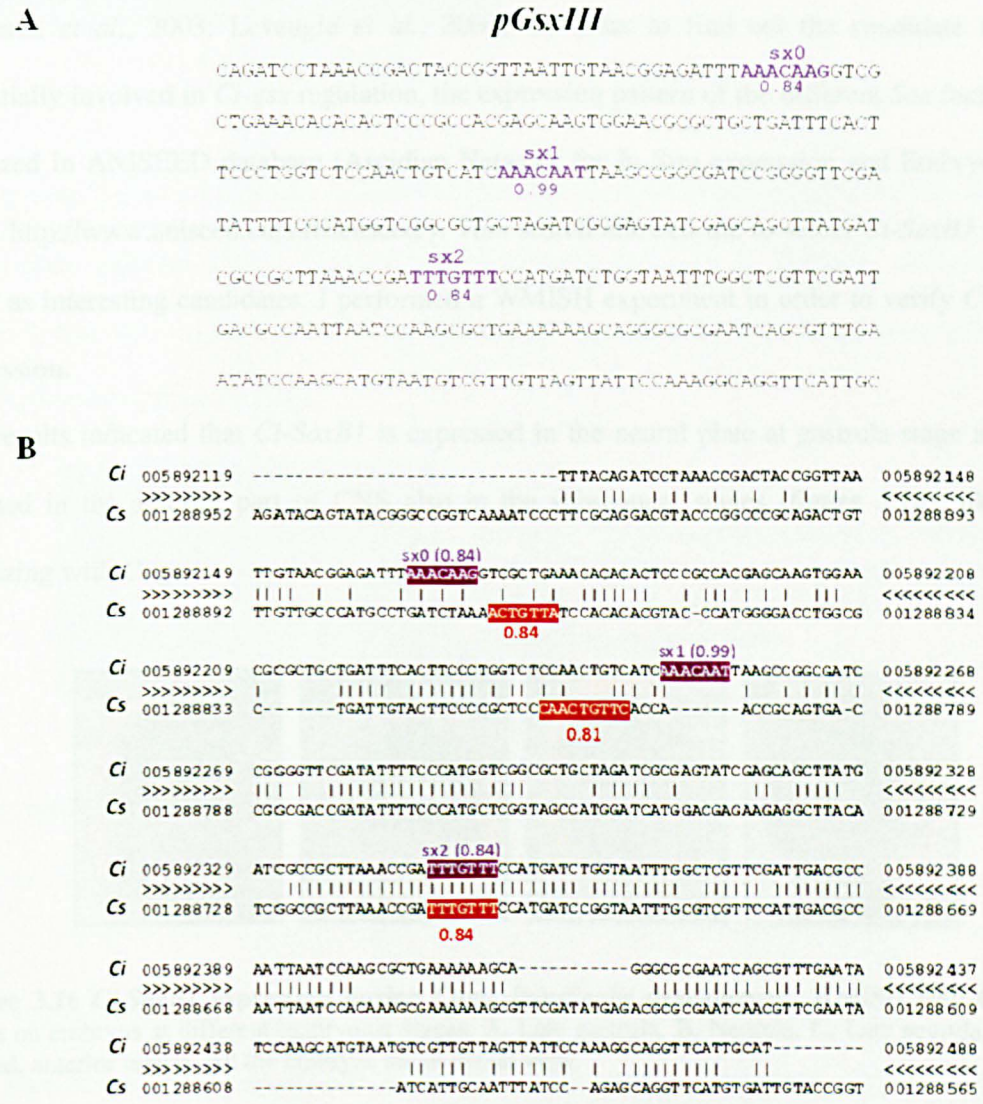


Figure 3.15 ***In silico*** **analysis of pGsxIII region.** Relative scores are indicated in correspondence of each predicted binding site. **A.** Localization of the three putative Sox binding sites identified in *pGsxIII*. **B.** Putative Sox binding sites on *pGsxIII* compared to the corresponding region in *C. savignyi*. The alignment of *pGsxIII* region compared to the corresponding region in *Ciona savignyi* is reported. Ci: *C. intestinalis*; Cs: *C. savignyi*. Putative Sox binding site in *Ciona intestinalis* promoter are indicated in purple; putative Sox binding sites in the corresponding region in *C. savignyi* are indicated in red.

I also compared the result obtained on pGsxIII fragment with the corresponding region in *Ciona savignyi* genome. I found that sx2 was conserved between the two species, whereas sx0 and sx1 did not perfectly match with *C. savignyi* promoter; however the bioinformatic analysis performed on *C. savignyi* promoter revealed putative Sox binding sites very close to the *C. intestinalis* ones (***figure 3.15 B***).

3.6.1 Ci-SoxB1 and Ci-SoxC expression pattern

Seven members of Sox transcription factor family are present in *Ciona intestinalis* genome (Yamada *et al.*, 2003; Leveugle *et al.*, 2004). In order to find out the candidate member potentially involved in *Ci-gsx* regulation, the expression pattern of the different *Sox* factors was analyzed in ANISEED database (Ascidian Network for In Situ expression and Embryological Data, http://www.aniseed.cnrs.fr/aniseed/). This search allowed me to select *Ci-SoxB1* and *Ci-SoxC* as interesting candidates. I performed a WMISH experiment in order to verify *Ci-SoxB1* expression.

The results indicated that *Ci-SoxB1* is expressed in the neural plate at gastrula stage and then retained in the anterior part of CNS also in the subsequent stages (***figure 3.16***), likely co-localizing with *Ci-gsx*.

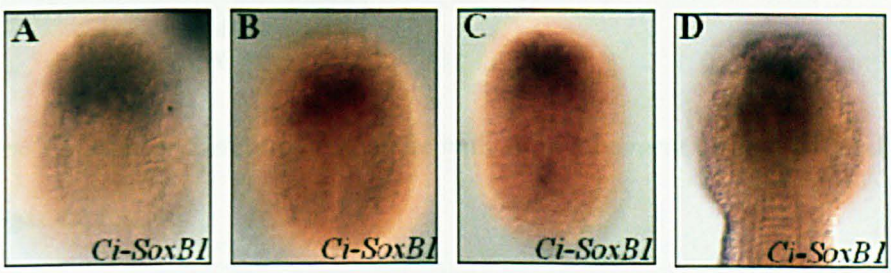


Figure 3.16 *Ci-SoxB1* expression during *Ciona intestinalis* development. WMISH with *Ci-SoxB1* probe on embryos at different embryonic stages. **A.** Late gastrula, **B.** Neurula, **C.** Late neurula, **D.** Late tailbud, anterior region. All the embryos are in dorsal view.

To assess this hypothesis I performed a double *in situ* hybridization experiment and the data confirmed that *Ci-gsx* and *Ci-SoxB1* overlaps each other in the territory where *Ci-gsx* is

expressed since early stages of development. Thus Ci-SoxB1 could represent a good candidate for *Ci-gsx* transcriptional regulation (***figure 3.17***).

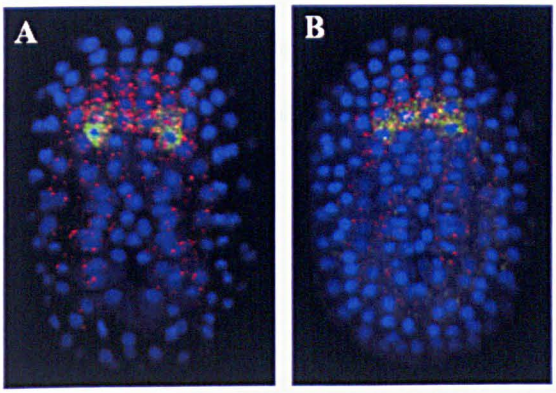


Figure 3.17 Positional relationship of Ci-SoxB1 and Ci-gsx in early developmental stages. Merge of *Ci-SoxB1* (in red), *Ci-gsx* (in green) and DAPI (blue, nuclear staining) is reported. **A.** Late gastrula stage, dorsal view. **B.** Neurula stage, dorsal view.

Concerning *Ci-SoxC*, I reported in ***figure 3.18*** some expression data from ANISEED database. These data, together with a personal communication from Dr C. Hudson, allowed me to conclude that also *Ci-SoxC* presents an interesting expression pattern, being detectable in the neural plate since gastrula stage.

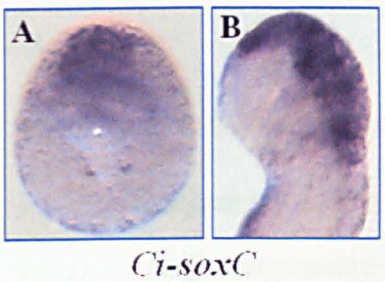


Figure 3.18 *Ci-SoxC* expression as reported in ANISEED database. A. Gastrula stage, dosarl view. **B.** Middle tailbud stage, anterior region, dorso-lateral view.

These expression data are therefore compatible with an involvement of both Ci-SoxB1 and Ci-SoxC in *Ci-gsx* expression.

3.6.2 Verifying potential Sox binding sites: further dissection of the promoter

Two approaches have been used in order to verify the role of the putative Sox binding sites found *in silico*. Firstly, I further dissected *pGsxIII* by PCR with suitable primers designed with the purpose of excluding one by one the three predicted sites; secondly I mutagenized these putative binding sites.

Concerning the first strategy, through the new dissections I obtained three constructs progressively shorter: *pGsx12>mCherry,* (314 bp), in which sx0 putative sites has been excluded; pGsx13>mCherry (232 bp), in which both sx0 and sx1 are lacking, and finally pGsx14>mCherry (136 bp), lacking any putative Sox binding site (***figure 3.19***).

The electroporation results have been assayed on embryos at larval stage and checked under fluorescent microscopy, in order to detect the possible alteration in the expression of the transgenic reporter.

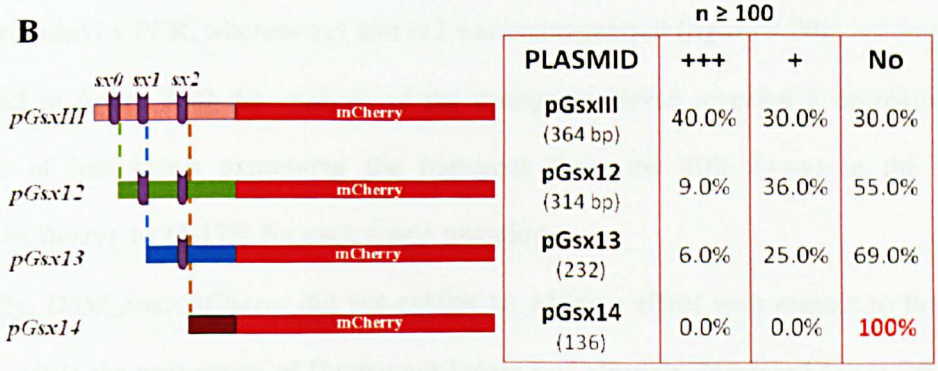


Figure 3.19 Dissection of *pGsxIII* regulatory region. A. Sequence of *pGsxIII* fragment, Putative sox binding sites are indicated in purple. Starting point of *pGsx12*, *pGsx13* and *pGsx14* are indicated in green, light blue and brown respectively. **B.** On the left: schematic representation of the analyzed fragment. Purple symbols indicate the position of sx0, sx1 and sx2. Table on the right: percentage of expression for each tested fragment, whose name and length is indicated in the first column ("Plasmid"). Column "+++" : percentages of larvae in which the expression of the reporter is more robust. Column "+": percentages of larvae in which the fluorescence is detectable, but the signal seems less bright. Column "No": percentage of larvae in which the fluorescence is not detectable. Note that pGsx14 is not able to activate the expression of the transgene.

As indicated by the result in ***figure 3.19*** the 50 bp regions affected by the first shortening is the most important one, since the percentage of larvae which were able to express the transgene dropped down from 40% to 9% (+++). In addition, the shortest fragment, pGsx14, which lacks any of the identified Sox binding sites, was not active at all (***figure 3.19 B***). This suggest that *sx0* site could be the stronger Sox element for *Ci-gsx* proper expression and that *sx1* and *sx2* collaborate with sx0 for the full expression of the transgene.

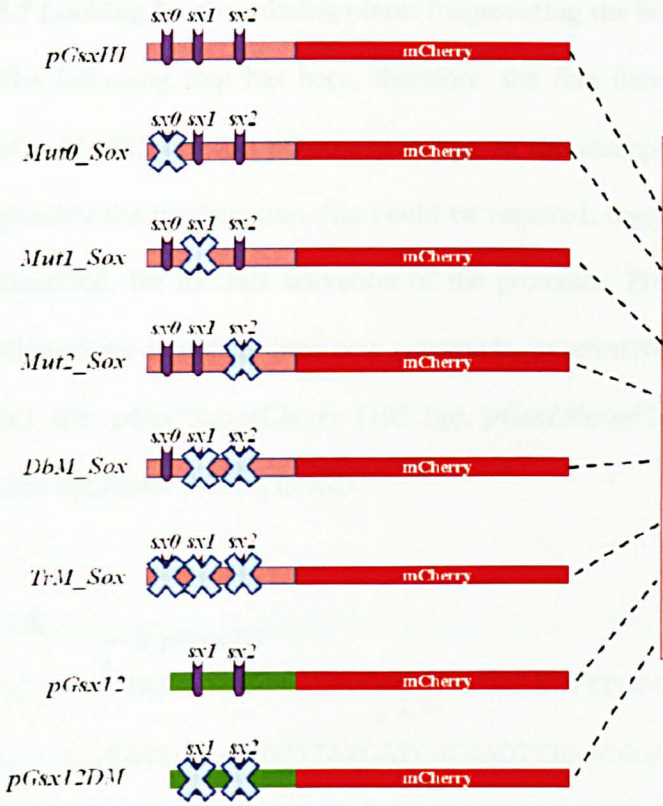
3.6.3 Verifying potential Sox binding sites: site-directed mutagenesis

As previously mentioned, the second approach consisted in the site-directed substitution mutagenesis of the putative Sox binding sites on *pGsxIII>mCherry* plasmid, totally changing the core of each one of them.

Hence, all three binding sites were mutated separately, generating the constructs *Mut0_Sox>mChe*, *Mut1_Sox>mChe*, *Mut2_Sox>mChe*; moreover, I combined the mutations in sx1 and sx2 sites in the construct *DbM_Sox>mChe*, while all the sites were mutated together in *TrM_Sox>mChe*. In parallel, I also prepared the plasmid *pGsx12DM>mCherry*, in which sx0 site was excluded by PCR, whereas sx1 and sx2 were mutagenized (***figure 3.20***).

As reported in ***figure 3.20*** the analysis of the transgenic larvae revealed a decrease in the percentage of individuals expressing the transgene from the 40% (+++) in the control (*pGsxIII>mCherry*) to 15-12% for each single mutation.

Surprisingly, *DbM_Sox>mCherry* did not exhibit an additive effect with respect to the single mutations, while the percentage of fluorescent larvae was strongly decreased (+++: 2%) when the 3 Sox sites were all mutated (*TrM_Sox>mChe*). The result of *pGsx12DM>mCherry* electroporated embryos almost perfectly paralleled the triple mutant, thus indicating that sx0 is probably the key site responsible for the strong decrease of promoter activity observed when comparing *pGsxIII* and *pGsx12* regions (***figure 3.20***).



n ≥ 100

PLASMID	+++	+	No
pGsxIII	40.0%	30.0%	30.0%
Mut0	15.0%	25.0%	60.0%
Mut1	13.0%	29.0%	58.0%
Mut2	12.0%	32.0%	56.0%
DbM	14.0%	26.0%	60.0%
TrM	2.0%	18.0%	80.0%
pGsx12	9.0%	36.0%	55.0%
pGsx12DM	0.0%	20.0%	80.0%

Figure 3.20 Scheme of the results obtained by site-directed mutagenesis of the three predicted Sox binding sites in pGsxIII region. On the **left** a graphical representation of each tested construct is reported. Purple symbols indicate the three putative Sox binding sites; a light blue cross indicates mutated binding sites. The table on the **right** shows percentage of expression for each tested plasmid. Column "+++" : percentages of larvae in which the expression of the reporter is more robust. Column "+": percentages of larvae in which the fluorescence is detectable, but the signal seems less bright. Column "No": percentage of larvae in which the fluorescence is not detectable.

Taken together these results suggest that the identified Sox binding sites are involved in the activation of *Ci-gsx*, acting in a co-operative mechanism that requires the presence of all of them.

Given the residual activity driven by the triple mutant (+++: 2%, +: 18%) and pGsx12DM (+: 20%), one can easily suppose the requirements of further player(s) in the *Ci-gsx* regulation. The data presented in ***figure 3.19 B*** and ***3.20*** indicate that the "missing player" could act through the binding on the region included between pGsx13 and pGsx14 fragments (that is included between sx1 and sx2).

3.7 Looking for the missing piece: fragmenting the fragment

The following step has been, therefore, the fine dissection of the region included between *pGsx13* (232 bp) and *pGsx14* (136 bp), in the attempt to identify the minimal promoter and possibly the binding sites that could be required, together with the Sox elements previously identified, for the full activation of the promoter. Primers specifically designed to this aim allowed me to obtain three new constructs, progressively shorter at the 5' end and containing sx2 site: *pGsx13a>mCherry* (195 bp), *pGsx13b>mCherry* (181 bp) and *pGsx13c>mCherry* (159 bp; ***figure 3.21 A***, in red).

A

B

PLASMID	+++	+	No
pGsxIII (364 bp)	40.0%	30.0%	30.0%
pGsx13 (232 bp)	6.0%	25.0%	69.0%
pGsx13a (195 bp)	4.0%	21.0%	75.0%
pGsx13a2 (185 bp)	3.0%	22.0%	75.0%
pGsx13b (181 bp)	0.0%	0.0%	**100%**
pGsx13c (160 bp)	0.0%	0.0%	**100%**

n≥100

Figure 3.21 Dissection of the region included between pGsx13 and pGsx14 fragments.
A. Sequence of *pGsx13* fragment. Sx2 site is indicated in purple. The starting point of the fragments pGsx13 and pGsx14, previously obtained, are indicated by green and brown arrows, respectively. The starting point of *pGsx13a*, *pGsx13b* and *pGsx13c* are indicated by red arrows. The starting point of the subsequent dissection, *pGsx13a1* is indicated in green. **B.** Percentage of expression for each tested fragment, whose name and length is indicated in the first column ("Plasmid"). Column "+++" : percentages of larvae in which the expression of the reporter is more robust. Column "+": percentages of larvae in which the fluorescence is detectable, but the signal seems less bright. Column "No": percentage of larvae in which the fluorescence is not detectable. Note that pGsx13b and pGsx13C are not able to activate the expression of the transgene.

Later, on the basis of the results obtained from the this dissection, I further subdivided the region included between pGsx13a and pGsx13b, thus generating the constructs *pGsx13a2>mCherry* (185 bp; ***figure 3.21 A***, in green).

Looking at the ***picture 3.21 B*** it seems clear that the region comprised in the few bases between *pGsx13a2* and *pGsx13b* is involved, together with Sox binding sites previously identified, in Ci-gsx activation.

It is also interesting to highlight that for no one of the active constructs I detected an ectopic expression of *mCherry* reporter gene.

3.8 Searching for putative transcription factor binding sites in pGsxIII regulatory region (part II)

In parallel with the JASPAR *in silico* analysis already reported, I also used Genomatix software to look for other interesting candidates possibly involved in *Ci-gsx* regulation. In particular, I focused my attention on two putative binding sites for Msx (***figure 3.22***, red rectangles).

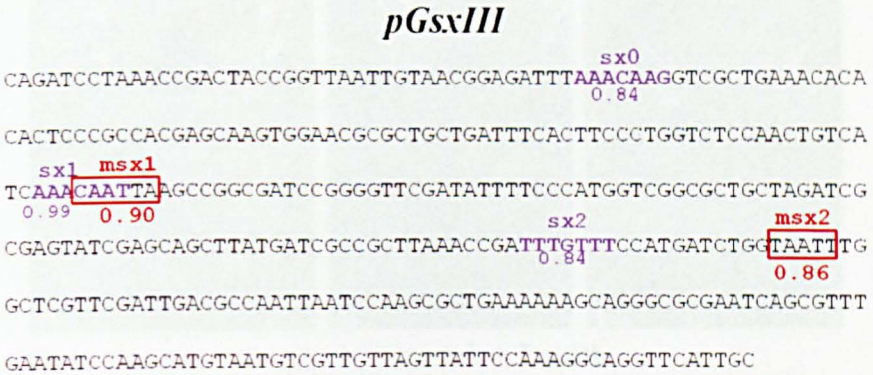


Figure 3.22 Localization of the two putative Msx binding sites *in silico* (Genomatix software) identified in *pGsxIII* region. Sox putative binding sites are indicated in purple, and Msx predicted binding sites are highlighted in red rectangles. Relative scores are reported below each predicted binding site.

Msx ortholog in *Drosophila* is directly regulated by the ortholog of Gsx in a repressive one-way cascade, fundamental in the DV patterning of trunk neuroectoderm. Conversely, in procephalic neuroectoderm regions Msx is also able to repress *Gsx*. In both cases, the two genes are expressed in adjacent, never overlapping territories (see Chapter 1).

One Msx ortholog has been found in Ciona (*Ci-msxb*); its expression pattern has been already characterized some years ago by our research group (Aniello *et al.*, 1999).

3.8.1 Comparison between Ci-gsx and Ci-msxb expression pattern

In order to investigate the putative relation between Ci-gsx and *Ci-msxb*, I compared their expression patterns by double *in situ* hybridization experiments.

The two genes are clearly expressed in adjacent territories, as occurs in *Drosophila*, since gastrula stage. At this stage *Ci-msxb* marks the pigment cells precursors (***figure 3.23***, white arrows) , which are in close proximity to *Ci-gsx*, but it is also present in a much wider territory, including mesenchyme cells and muscle precursors.

Given the *in situ* results, I started few preliminary experiments to test any repressor activity between the two factors.

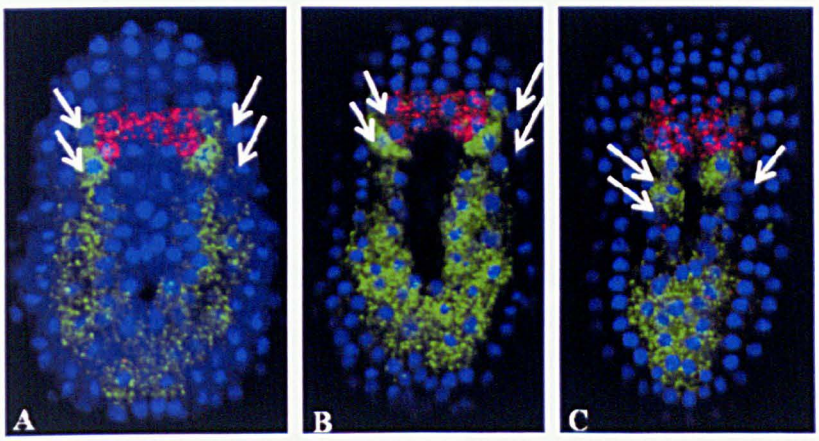


Figure 3.23 Positional relationship of *Ci-msxb* and *Ci-gsx*. Merge of *Ci-msxb* (in green), *Ci-gsx* (in red) and DAPI (blue, nuclear staining) is reported. White arrows indicate the pigment cells lineage. **A.** Late gastrula stage, dorsal view. **B.** Neurula stage, dorsal view. **C.** Late neurula stage, dorsal view.

3.8.2 Ectopic expression of Ci-msxb: effects on Ci-gsx

In order to test if Ci-msxb is able to repress *Ci-gsx*, I expressed ectopically Ci-msxb in Ci-gsx territories by exploiting *pGsx* promoter.

Embryos at neurula stage electroporated with *pGsx>Msx* were analyzed through *in situ* hybridization experiments, in order to check the expression of *Ci-gsx*. Embryos electroporated with *pGsx>mCherry* were used as a control.

The ***picture 3.24*** (***A, B***) clearly shows that *Ci-gsx* signal was strongly reduced or totally absent in most of the observed embryos.

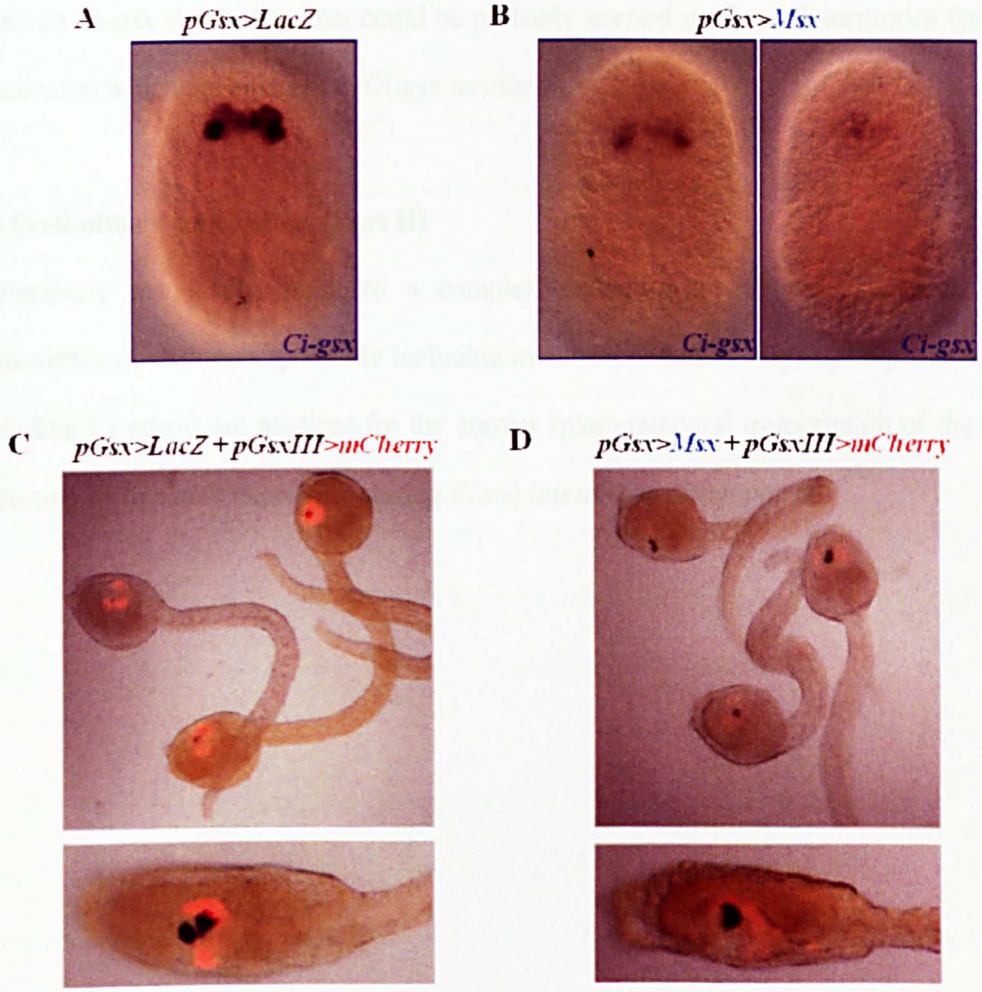


Figure 3.24 Ectopic expression of *Ci-msxb in Ci-gsx* territories. A, B. WMISH with *Ci-gsx* probe on control embryos at neurula stage electroporated with *pGsx>LacZ* (A) and *pGsx>Msx* (B). Dorsal view, anterior is on the top. The experiment has been performed in three different batches of animals, and for each condition at least 25 embryos were analyzed.
C, D. mCherry fluorescence detected in late tailbud III embryos (on the top) and larvae (at the bottom) electroporated with (C) *pGsx>LacZ* + *pGsxIII>mCherry* (control samples) and (D) *pGsx>Msx* + *pGsxIII>mCherry*. The experiment has been performed in three different batches of animals, and for each condition at least 100 embryos were analyzed.

Hence, I tested the capability of Ci-msxb to interfere with the function of *pGsxIII* promoter, with an experiment similar to the one reported in paragraph 3.4.

The construct *pGsx>Msx* was co-electroporated with *pGsxIII>mCherry* and the putative alteration of mCherry expression were checked in embryos at the late tailbud and larval stages.

Late tailbud/larvae co-electroporated with *pGsx>Lacz* + *pGsxIII>mCherry* were used as control.

The data clearly indicated that the activity of *pGsxIII* promoter was strongly decreased in *pGsx>Msx* electroporated embryos (***figure 3.24 C, D***).

The preliminary data so far described support the hypothesis of a possible negative effect of Ci-msxb on Ci-gsx expression, that could be probably exerted in Ci-msxb territories (pigment cell precursors) which are adjacent to Ci-gsx territories.

3.9 Preliminary conclusions (Part II)

Collectively my results point to a complex regulation of *Ci-gsx* expression, in which transcriptional activators (probably including members of Sox family) and repressors (probably including Ci-msxb) are required for the correct spatio-temporal transcription of the gene, thus reflecting its dynamic expression during *Ciona intestinalis* development.

Chapter 4

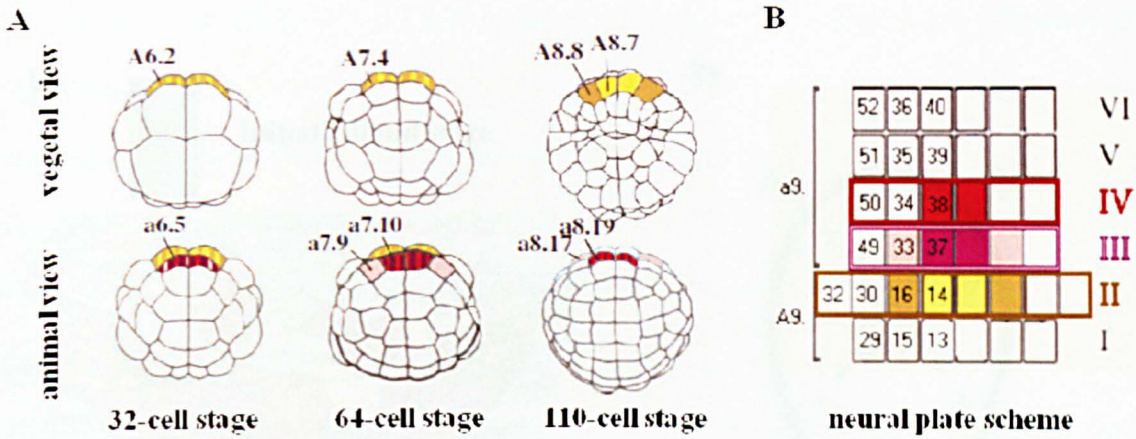
DISCUSSION

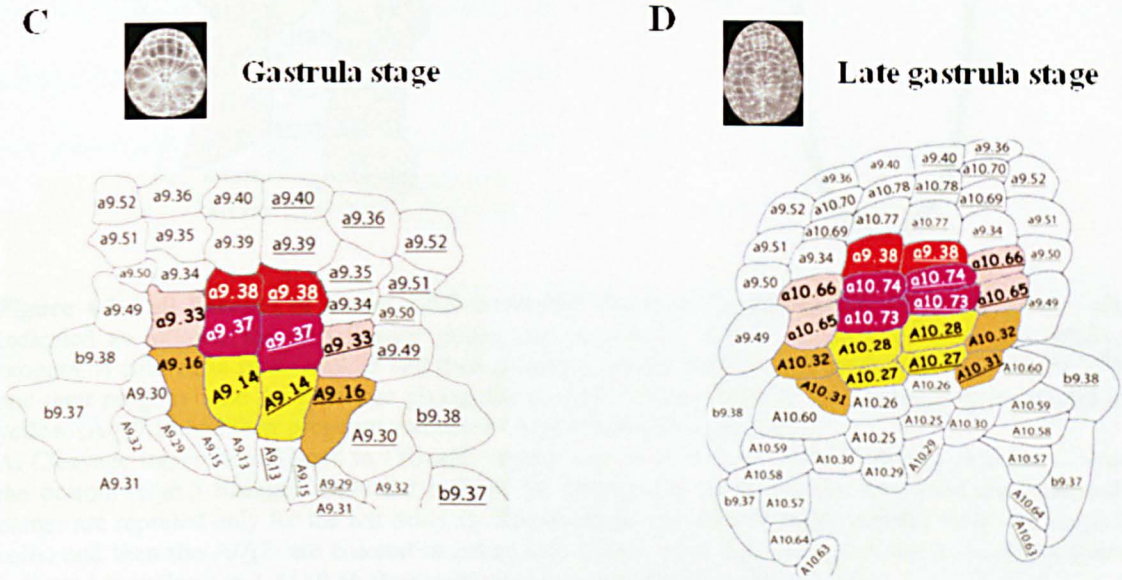
From its first identification in mouse (Singh *et al.*, 1991; Hsieh-Li *et al.*, 1995; Valerius *et al.*, 1995), and then in medaka fish (Deschet *et al.*, 1998), amphioxus (Brooke *et al.*, 1998) and *Drosophila* (Weiss *et al.*, 1998), Gsx has been found almost in all living animals, from placozoans and cnidarians up to vertebrates. Although being so "ancient", *Gsx* gene did not undergo many duplications during its evolution, with the only exception of the best known event occurred in vertebrate lineage, the whole genome duplication, which gave rise to the paralogs *Gsh1* and *Gsh2*. Expression and functional analyses of this transcription factor in all the species so far investigated indicate that its principal function is related to nervous system formation, so that the study of Gsx results particularly interesting in order to get insights into the early evolution of eumetazoan nervous system development during embryogenesis.

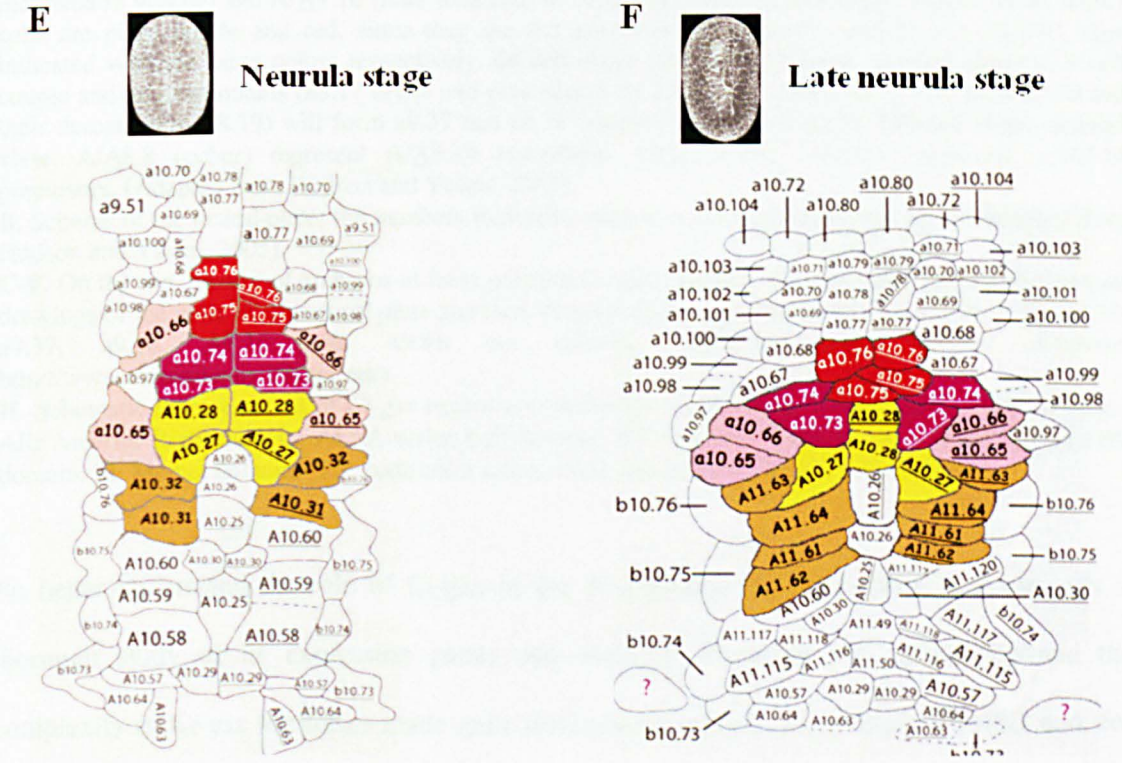
As hereinafter discussed, these evolutionary aspects intersect the main interest of our research group, which focuses on the determination of the gene regulatory network underlying the formation of pigmented sensory organs in the ascidian *Ciona intestinalis*.

4.1 *Ci-gsx* in the developing *Ciona intestinalis* Central Nervous System

When I started my PhD, *Ci-gsx* was selected as a putative candidate for the early specification of photoreceptor cells included in *Ciona* ocellus. At that time, literature data about *Ci-gsx* were quite poor. Concerning its expression, *Ci-gsx* was reported to be present in the a9.33 cell couple and then in their descendants (a10.65 and a10.66 cell pairs) during gastrula/late gastrula stages (Hudson and Lemaire, 2001). According to this work, neither the a9.37 cells (two cells located between the a9.33) neither their progeny (a10.74 and a10.73 couples) resulted stained. Later on in development, *Ci-gsx* was described in two domains bordering the midline of the embryo and then in the posterior sensory vesicle at tailbud/late tailbud stages. This description did not cover developmental times between late gastrula and tailbud stages; furthermore, the extent of the reported territories at tailbud stages could not be framed as progeny of the sole a9.33 cell lineage.


A
vegetal view
animal view
A6.2
A7.4
A8.8
A8.7
a6.5
a7.9
a7.10
a8.17
a8.19
32-cell stage
64-cell stage
110-cell stage
B
52 36 40
51 35 39
50 34 38
49 33 37
32 30 16 14
29 15 13
a9.
A9.
VI
V
IV
III
II
I
neural plate scheme


C
Gastrula stage
D
Late gastrula stage


E
Neurula stage
F
Late neurula stage

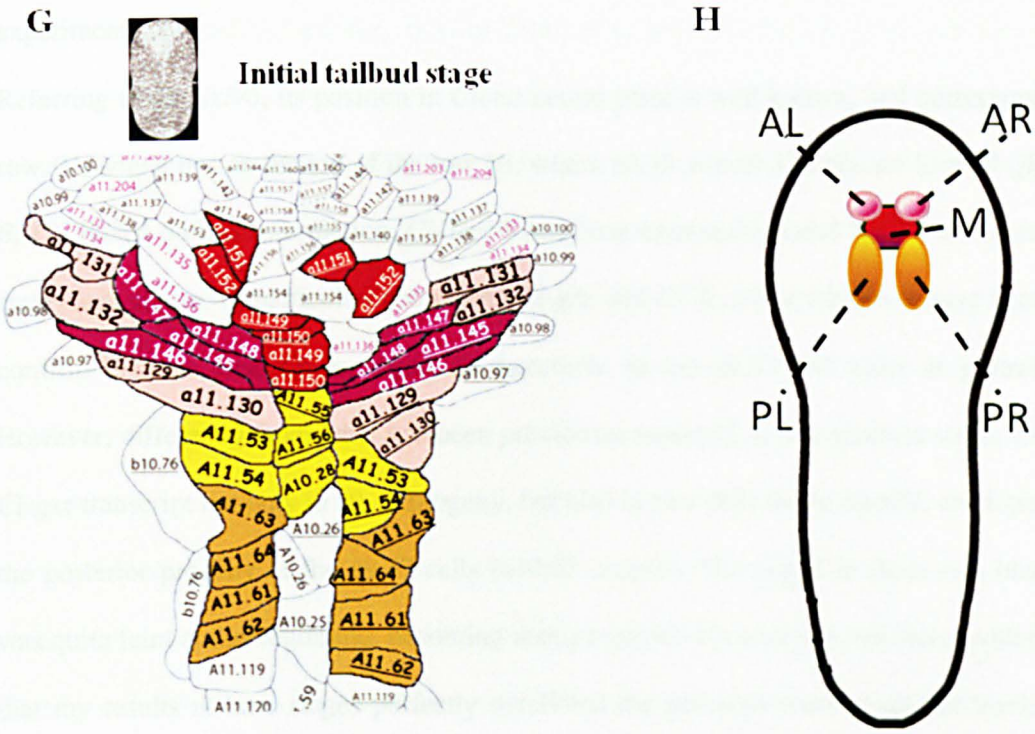


Figure 4.1 Cell lineages of Ciona CNS territories discussed in paragraph 4.1. Cell lineages are indicated as follows: the a6.5 lineage giving rise to a/a9.33, a/a9.37 and a/a9.38 cells and relative progeny is colored in pink (a/a9.33 and their progeny), purple (a/a9.37 and their progeny), red (a/a9.38 and their progeny); the A6.2 lineage giving rise to A/A9.14 and A/A9.16 and its progeny is colored in yellow (A/A9.14 and their progeny) and ocher (A/A9.16 and their progeny).
A. Cleavage stages from 32-cell to 110-cell; vegetal view is on the top (A/A6.2 lineage), animal view at the bottom (a/a6.5 lineage). Only the cells of the lineages in which we are interested are indicated; names are reported only for the left progeny. **32-cell stage and 64-cell stage, animal view**: the A/A6.2 cells, and then the A/A7. are colored in ocher and yellow since they will give rise to A/A9.14 (later indicated in yellow) and A/A9.16 (later indicated in ocher) lineages. **32-cell stage, vegetal view**: a/a6.5 cells are pink, purple and red, since they are the precursors of a/a9.33, a/a9.37 and a/a9.38, later indicated with the same colors respectively. **64-cell stage and 110-cell stage, vegetal view:** a7.9 cell couple and its descendants (a8.17 cells) will give rise to a9.33 lineage (pink color). The cells a7.10 and their descendants (a8.19) will form a9.37 and a9.38 couples (purple and pink). **110-cell stage, animal view**: A/A8.8 (ocher) represent A/A9.16 precursors, while A/A8.7 (yellow) represent A/A9.14 precursors. (Adapted from Hudson and Yasuo, 2005).
B. Scheme of the neural plate; the numbers indicating each row are reported on the right. (Adapted from Hudson and Yasuo, 2005).
C-F. On the top, picture of embryos at from gastrula to initial tailbud stage are reported. At the bottom, drawings of the cells of the neural plate and their descendants are indicated. Only the cells of the a9.33, a9.37, a9.38, A9.14 and A9.16 are colored. (Adapted from Aniseed database, http://www.aniseed.cnrs.fr/aniseed/).
H. Schematic representation of *Ci-gsx* expression territories compared to *Ci-Meis* ones at tailbud stage. AR: Anterior Right domain; AL: Anterior Left domain; PR: Posterior Right domain; PL: Posterior Left domain; M: Medial domain. The code color refers to the one used in the previous pictures.

To better understand the role of Ci-gsx in the development of Ciona photoreceptor cells, a thorough study of its expression patter was required. However, the dynamicity and the complexity of *Ci-gsx* territories made quite difficult the analysis from single WMISH and, for

this reason, I used *Ci-Six3/6* and *Ci-Meis* probes as landmarks in double *in situ* hybridization experiments.

Referring to *Ci-Six3/6*, its position in *Ciona* neural plate is well known, and corresponds to the row IV, which lies on the top of the row III, where a9.33 and a9.37 cells are located (***figure 4.1 B, C***). Based on this information, *Ci-Six3/6* has been extremely useful for the analyses at early embryonic stages. In particular, comparing *Ci-gsx* and *Ci-Six3/6* territories, I have been able to confirm that the first *Ci-gsx* signal is detectable in the a9.33 cell pairs at gastrula stage. However, differently from what has been previously reported, at late gastrula stage, I observed *Ci-gsx* transcript not only in a9.33 progeny, but also in two cells in the middle, corresponding to the posterior progeny of the a9.37 cells (a10.73 couple). The signal in these two blastomeres was quite faint at the beginning, becoming then progressively stronger and more evident. Given that my results at later stages perfectly paralleled the previous work, I do not think that the observed discrepancy has a real "biological" meaning; probably it was simply due to a slight difference in the embryonic stage, or to a greater sensitivity of the *in situ* protocol that we presently use (Christiaen *et al.*, 2009), with respect to the one used in Hudson and Lemaire (2001; Wada *et al.*, 1995) .

For what concerns *Ci-Meis* instead, it has been useful for comparisons with *Ci-gsx* from late gastrula to late neurula stages. Clear data were already available from middle-late tailbud stages, when *Ci-Meis* is expressed, together with *Ci-TH* (a gene that codes for a key enzyme in dopamine biosynthesis), in a small population of DA-synthetizing cells in the ventromedial portion of *Ciona* sensory vesicle, included in the coronet cells (Moret *et al.*, 2005a). Conversely, the *Ci-Meis* early expression was not previously analyzed, and it was only hypothesized that *Ci-Meis* positive cells at the tailbud and larval stages were derived from the IV row of the neural plate, in which also *Ci-Six3/6* is expressed, and in particular from the a9.34 or a9.38 cell couples (Moret *et al.*, 2005b; ***figure 4.1 B, C***).

My results at early embryonic stages did not fit with the previous hypothesis of Moret and co-workers (2005b), since I found that *Ci-Meis* at late gastrula stage was expressed only in the four a9.37 descendants (a10.73 and a10.74 pairs), so more posteriorly to what has been proposed in

the past (***figure 4.1 D***, purple cells). The signal was retained in these four cells up to late neurula stage; after this stage, it marked in total eight cells, that could correspond to the progeny of a10.73 and a 10.74 couples (***figure 4.1 G***, purple cells, compared to ***figure 3.5 L, N***). Alternatively, the posterior four cells of *Ci-Meis* domain at this stage could represent the a10.73 and a10.74 couples, not yet divided (***figure 4.1 E-F***, purple cells), while the anterior four cells could correspond to the derivatives of a9.38 cell pair (a10.75 and a10.76 couples, ***figure 4.1 E-F***, red cells). The latter hypothesis would be in agreement with the previous theory of Moret and co-workers (2005b). Since these changes occur quite rapidly, and the difference between late neurula/initial tailbud is really subtle, it is difficult to say if a10.73 and a10.74 cell pairs are already divided when *Ci-Meis* marks eight cells, so, with the data obtained so far, I am not able to draw a certain conclusion. According to Cole and Meinertzhagen (2004), the a10.73 and a 10.74 cells undergo a new mitotic division only at early tailbud stage (which follows initial tailbud stage); in addition this division occurs along the dorso-ventral axis, and not aterior-posteriorly. These two observations would suggest that at late neurula/initial tailbud stage Ci-Meis territory could correspond both to a9.37 and a9.38 lineages.

From early tailbud stage, instead, it was difficult to follow *Ci-Meis* expressing cells, since its domain sank in the head, making difficult the observation at the confocal microscope.

The *Ci-Meis* expression pattern compared to *Ci-gsx* indicated that *Ci-gsx* co-localizes with *Ci-Meis* in the posterior progeny of the a9.37 cells at late gastrula stage, and later (up to neurula stage) it expanded also in the anterior a9.37 descendants. I could, therefore, conclude that at neurula stage *Ci-gsx* was expressed in eight cells corresponding to the a9.33 and a9.37 derivatives.

From late neurula stage *Ci-gsx* gradually faded out in the *Ci-Meis* overlapping territories and, at middle/late tailbud stage, the expression of the two genes became mutually exclusive, thus indicating that, at this stage, *Ci-gsx* likely disappeared from the a9.37 progeny.

What is intriguing was the repositioning of *Ci-gsx* expression territories between late-neurula/tailbud stage that led to a progressive establishment of two distinct symmetric domains

along the anterior posterior axis. The anterior one consisted in a circumscribed area, whereas the posterior one was wider and appeared as bilateral stripes (***figure 4.1 H***).

Based on a careful analysis of the mitotic divisions of the cells and the alignment of their progeny along the neural tube (based on Cole and Meinertzhagen, 2004 in combination with ANISEED data), one could suppose that the anterior bilateral domains might correspond to the progeny (whole or partial) of a9.33 and (possibly) a9.37 cells, which mainly undergo dorso-ventral oriented mitotic divisions since early tailbud stage (***figure 4.1 H***, AL and AR spots).

Furthermore, the co-localization of *Ci-gsx* with *Ci-Rx*, at tailbud stage, let easily argue that the anterior right spot corresponded to photoreceptor cell precursors. This data appeared to be in agreement with previous studies which indicated that photoreceptor cells of the ocellus derive from the right progeny of a9.33 and possibly a9.37 cells (Cole and Meinertzhagen, 2004; Horie *et al.*, 2005).

While *Ci-gsx* endogenous transcript was not detectable at larval stage, the mCherry reporter protein, being very stable, remained longer visible in the *Ci-gsx* lineages. The analysis of the larvae expressing *pGsx>mCherry* transgene allowed to argue that the anterior left *Ci-gsx* spot, together with the *Ci-gsx/Ci-Meis* "medial" territories (***figure 4.1 H***, AL and M spots), became part of the ventral-posterior sensory vesicle. Most of the electroporated larvae presented, indeed, two fluorescent stained areas: one in correspondence of ocellus photoreceptor cells and another one localized close to the otolith (***figure 4.2***).

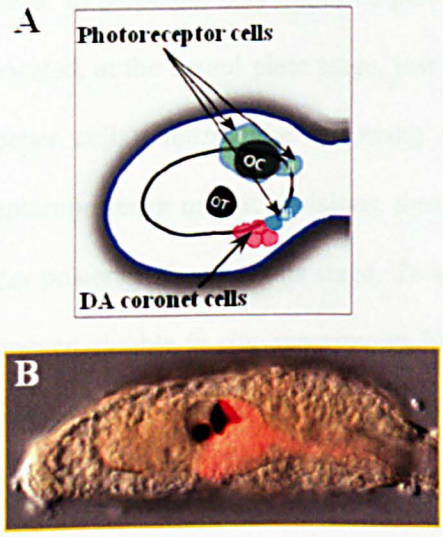


Figure 4.2 The expression of mCherry reporter in a transgenic larva electroporated with *pGsx>mChe* could include DA cells and group III photoreceptors.
A. Schematic representation of larval sensory vesicle. Otolith (OT) and Ocellus (OC) are grossly shown in black. The light blue spots indicate the three groups of photoreceptor cells (PRCs) present in the sensory vesicle. The coronet cells, in close relationship with the population III of PRCs, are indicated in pink (Adapted from Razy-Krajka *et al.*, 2012).
B. Trunk of a transgenic larva electroporated with *pGsx>mCherry*.

It is noteworthy that the ventral part of the sensory vesicle includes the dopaminergic cells, delineated by *Ci-Meis* expression, and the class III of photoreceptor cells. Do the *pGsx>mCherry* positive cells in larvae correspond to one or both territories?

A series of double electroporation, already scheduled, will enormously contribute to gain insights into this topic. In particular the *pCiTH>LacZ* construct, which specifically drives LacZ reporter in the dopaminergic cells of the larvae, will be instrumental, if electroporated in combination with *pGsx>mCherry*, to reveal any overlapping territories of LacZ/mCherry expression in the sensory vesicle. On the other hand, the *pArrestin>GFP* construct labels both the ocellus photoreceptors (class I and II) and class III of photoreceptors. Thus the *pArrestin>GFP* construct, used in combination with *pGsx>mCherry*, will permit to highlight if the reporters superimpose in both ocellus photoreceptors and photoreceptor class III close to the otolith.

This is what concerns the anterior territories of *Ci-gsx* expression. As previously mentioned, from the early tailbud stage *Ci-gsx* expression resolves in a posterior domain, which gradually expanded in two stripes (***figure 4.1 H***, PL and PR spots). Based on 1) the well described cell lineages of Ciona CNS; 2) the structure of the neural plate stages, when cells are precisely aligned in a grid-like organisation and each cell can be identified (Nishida, 1987; Nicol and Meinertzhagen, 1988b; Cole and Meinertzhagen, 2004); 3) the position of these cells in the tailbud, it was possible to suppose that the Ci-gsx posterior domains derived from the capital A cells. In particular they could be part of the progeny of A9.16 +/- A9.14 pairs. These cells are located, at the neural plate stage, just behind the a9.33/a9.37 pairs (***figure 4.1* B, C**, yellow and ocher cells); during the following steps of development, they alternate dorso-ventral and anteroposterior mitotic divisions, thus being compatible with the width and the extension of *Ci-gsx* posterior signal at this stage. To shed light on these posterior domains and reconstruct their lineage, double *in situ* experiments have been already planned. To this end two markers of the row II of the neural plate (Hudson *et al.*, 2007; Imai *et al.*, 2006) have been identified, *cicl007j15* clone and *Ci-FoxB* gene. For *cicl007j15* clone the sole data available on ANISEED indicates that it is present in A8.7 and A8.8 cell pairs at the early gastrula stage (***figure 4.1 A***,

yellow and ocher cells). *Ci-FoxB* pattern has been instead pretty well defined and, according to ANISEED, the gene is expressed in the row II of the neural plate at the late gastrula stage (***figure 4.3 A***); from late neurula up to the tailbud stage the expression is localized in two lateral bands in the presumptive posterior sensory vesicle, recalling *Ci-gsx* posterior territories (***figure 4.3 B*** and ***C*** compared to ***D***).

Thus these *Ci-FoxB* and *cicl007j15* genes, each in combination with *Ci-gsx*, will allow to delineate the fate of the posterior territories in which *Ci-gsx* is expressed.

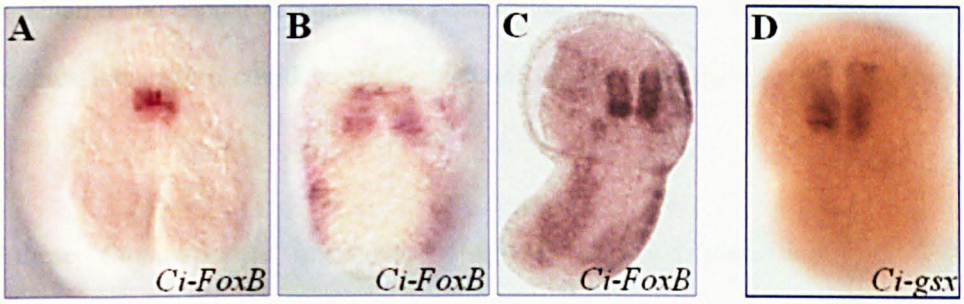


Figure 4.3 *Ci-FoxB* expression pattern, reported in ANISEED database, compared to *Ci-gsx*. A. Late gastrula stage embryo, dorsal view. Ci-FoxB is expressed in A10.27 and A10.28 cells. **B.** Late neurula stage, dorsal view. **C.** Middle tailbud stage, dorso-lateral view. **D.** WMISH with Ci-gsx probe on embryo at middle tailbud stage, dorsal view-

The precise definition of the fate of these cells is of particular interest in light of recent data (still not published) shown at the 7th Tunicate Meeting hold in Naples. During the poster session, the group of Dr Sasakura (Tanaka *et al.*), using DiI-labeling technique of single blastomeres at 32-cell stage, showed data indicating that both, photoreceptor cells of the ocellus and group III photoreceptor cells, could derive from the capital A lineage (right and left blastomere A6.2 lineages, respectively, of the 32 cell stage embryo; ***figure 4.1 A***, yellow/ocher cells). These results is in contrast with the previous hypothesis that assigned ocellus photoreceptor cell lineage to the right a6.5 blastomeres (progenitor of the right a9.33 and a9.37 cells; ***figure 4.1 A***, pink/purple/red cells). Provided that these data are preliminary and need to be further investigated, it is undeniable that our findings on the anterior expression of *Ci-gsx* which coincides, at tailbud stage, with that of *Ci-Rx* (especially on the right side of the embryos), coupled with the already known function played by Ci-Rx in *Ciona* ocellus

differentiation (D'Aniello *et al.* 2006), strongly point to the small a lineage derivatives as being part of photoreceptor ocellus formation.

One possible speculation to explain these discrepancies could be that both, the small a and capital A derivatives, contribute to the formation of the whole set of photoreceptor cells of the ocellus. Actually, as already mentioned, in the ocellus organ are present two different types of photoreceptors, group I and group II, which show distinct morphologically features; thus we cannot exclude that each class of photoreceptors are generated by two different developmental lineages, small a and capital A. Under this perspective the detailed and comparative analysis of *Ci-gsx* expression pattern could be really instrumental to gain more insights on this intricate story.

Beyond these "*Ciona*-specific" questions, the interest for *Ci-gsx* expression in the CNS is also linked to some intriguing evolutionary interrogatives about the relationship between ascidian sensory vesicle and vertebrate hypothalamus plus photoreceptive organs (Walls, 1942; Razy-Krajka *et al.*, 2012). Indeed the dopaminergic/coronet cells of *Ciona* are located in the expression domain of homologues of vertebrate hypothalamic markers. As dopaminergic/coronet cells are specified in the hypothalamus in all vertebrates, it has been speculated that ascidian ventral sensory vesicle is the remnant of a proto-hypothalamus that may have been present in the chordate ancestor (Moret *et al.*, 2005a; Moret *et al.*, 2005b). On the other hand, the dopaminergic/coronet cells show also many molecular and functional characteristics of amacrine cells, which are auxiliary cells present in the vertebrate retina (Razy-Krajka *et al.*, 2012). On these grounds, it has been proposed that the dopamine/coronet cells of the ascidian larva derive from an ancestral multifunctional cell population located in the periventricular, photoreceptive field of the anterior neural tube of chordates, which also gives rise to both anterior hypothalamus and the retina in craniates/vertebrates (Razy-Krajka *et al.*, 2012). From this perspective, the definition of *Ci-gsx* expression domains, that seem to cover both, the three photoreceptor types and the dopaminergic/coronet cells, will undoubtedly give important cues on the evolutionary relationships among these structures.

A more general consideration, still in light of my data, is that one of the most striking feature of Gsx in animal kingdom pertains the complexity of its expression pattern, that turns out to be very dynamic during embryonic development and often widespread in the CNS. Among deuterostomes, Gsx territories include numerous areas of forebrain, midbrain, hindbrain and spinal cord in vertebrates (Hsieh-Li *et al.*, 1995; Valerius *et al.*, 1995; Deschet *et al.*, 1998; Cheesman and Eisen, 2004; Illes *et al.*, 2009; ***figure 4.4 B-D***), while in other *taxa*, the presence of the gene is restricted to a few cells generally belonging to ectodermal/neural territories, as occurs for instance in amphioxus (Brooke *et al.*, 1998; Osborne *et al.*, 2009; ***figure 4.4 E-F'***) or sea urchin (Arnone *et al.*, 2006). Among protostomes, *Gsx* expression has been documented in the insects *Drosophila* and *Tribolium*, and the annelida *Capitella* and *Nereis virens*. Insect *Ind(Gsx)* domains cover both procephalic and trunk neuroectoderm regions. In these organisms, Gsx is generally expressed quite early during embryogenesis, with a typical bilateral symmetric distribution along the anterior-posterior axis (***figure 4.4 A***). In contrast, in the annelid *Capitella* Gsx is transiently expressed during early stages of brain formation into a small domain close to the anterior end of the CNS and is not observed in any other tissue, or at any other larval or juvenile stage. This is very different to the spatially and temporally dynamic expression of Gsx in other annelida such as *Nereis virens* (Kulakowa *et al.*, 2008) and *Platynereis dumerilii* (Hui *et al.*, 2009). In *Platynereis Gsx* expression is seen both in the developing apical organ and cerebral ganglia and, later, in the ventral plate of the trunk CNS. Besides the expression domains in prospective neural tissue, *Gsx* transcripts are also detected in the stomodeum and, at 6 days of larval development, *Gsx* expression is most prominent in small cell clusters of the cellularised gut, both in the region of the midgut and the posterior foregut (Kulakova *et al.*, 2008; Hui *et al.*, 2009). The expression in endodermal/gut region has been detected also in mollusks (Samadi and Steiner, 2010) and, for Gsh2, in *Xenopus tropicalis* (Illes *et al.*, 2009; ***figure 4.4 D***, red arrow).

Comparisons between *Platynereis* gene expression and the orthologs of deuterostomes lead to hypothesize that in the ancestor of eubilaterians *Gsx* expression pattern was complex, with *Gsx* domains in a variety of nervous system roles (eyes, neurosecretory cells and 'hindbrain', and

potentially along the anterior-posterior axis of the nerve cord). Gsx expression was then secondarily simplified in several lineages as in *Capitella*, amphioxus and possibly sea urchin. However, to distinguish whether the *Platynereis*–vertebrate comparison really does provide a better reflection of the ancestral condition than these simplified lineages (*Capitella*, amphioxus, sea urchin) requires a consensus regarding whether Gsx expression in the CNS is simple and anterior or complex and extended.

The expression pattern of *Ci-gsx* in our model organism can be placed in between the two mentioned opposite conditions. In *Ciona*, indeed, as described for most bilaterians, *Ci-gsx* is expressed quite early, with a dynamic bilateral pattern, in relation to which we observe variations of *Ci-gsx* positive domains, which remains symmetrical distributed along the anterior-posterior axis of the embryo (***figure 4.4 G***).

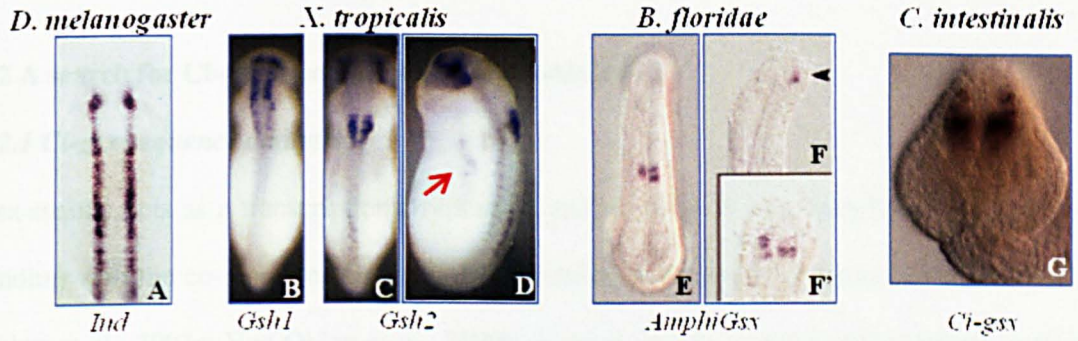


Figure 4.4 Examples of *Gsx* expression pattern in bilaterians. A. *Ind* expression in *Drosophila melanogaster*, stage 9 embryos, ventral view (Weiss *et al.*, 1998). **B-D.** *Gsh1* (B) and *Gsh2* (C, D) expression in *Xenopus tropicalis*, early tailbud stage 25 embryos. B, C: dorsal view; D: lateral view. The red arrow in D indicate Gsh2 expression lateral endoderm. (Adapted from Illes *et al.*, 2009). **E-F'.** *AmphiGsx* expression in Branchiostoma floridae. E: neurula embryo, were *AmphiGsx* is expressed in the neural tube. Dorsal view. F: late embryo, lateral view. *AmphiGsx* is expressed in the cerebral vesicle. F': magnified dorsal view of the embryo presented in F at the level of the arrowhead. (Osborne *et al.*, 2009). **G.** *Ci-gsx* expression in *Ciona intestinalis*, late tailbud embruo (anterior region), dorsal view.

Thus, *Ci-gsx*, differently from amphioxus or sea urchin, is not present only in a few "neural" cells; although expressed in a wider area, *Ci-gsx* signal is not detected in different territories along the whole CNS, as in most vertebrates, but it remains confined within the sensory vesicle precursors (homolog to the vertebrate fore-midbrain), thus being more reminiscent of what has been described in amphioxus (Brooke *et al.*, 1998; Osborne *et al.*, 2009; ***figure 4.4*** G compared

to ***A-F'***). Yet, compared to the amphioxus, *Ci-gsx* lineage includes different cell types, thus suggesting an involvement in multiple molecular pathways.

Provided that more species need to be examined in order to obtain a clearer picture about the expression of *Gsx*, it is tempting to speculate, based on our observations in *Ciona*, that the involvement of Gsx in different area of the whole CNS (head plus trunk) could be related to the complexity of the CNS of a given specie. In *Ciona* the trunk CNS is devoid of neuronal cell bodies and thus is not very specialized and functionally subdivided (Nicol and Meinertzhagen , 1991); in the trunk CNS *Ci-gsx* is not expressed. Most of the neuron soma are localized in the brain vesicle, where *Ci-gsx* is, actually, localized. One can thus hypothesize that, in the course of evolution and in the different taxa, *Gsx* expression territories expanded from the anterior to the posterior CNS in the lineages developing more complex neural structures in the trunk, like arthropods, annelids and higher chordates.

4.2 A search for Ci-gsx function in Ciona intestinalis

4.2.1 Ci-gsx sequence analysis

Gsx mainly acts as a transcriptional repressor, and it seems to exert this function through the binding with the co-repressor Groucho. This interaction occurs also, but not exclusively (Von Ohlen *et al.*, 2007a; Von Ohlen *et al.*, 2009), at the level of a conserved N-terminal motif, the Eh1-like domain (Engrailed homology 1), firstly identified in the homeodomain-containing proteins Engrailed (to which it owes its name) and Goosecoid, and then found also in other classes of transcription factors (Jiménez *et al.*, 1997; Jiménez et *al.*, 1999; Goldestein *et al.*, 2005).

Generally, the motif F-x-I-x-x-I (where x is any amino acid) is considered the typical minimum consensus for this domain, even if the alignment of various Eh1-like sequences reveals a substantial degeneracy with respect to this consensus. Indeed, a few amino acidic substitutions are already known, such as isoleucine (Ile, I) at position 3 or 6 replaced with valine (Val, V) or leucine (Leu, L) (Goldstein *et al.*, 2005), so that it would be probably more correct to assume the sequence F-x-B-x-x-B as consensus motif, where B is a branched hydrophobic residue (I, V

or L) (Dalafave, 2009). Besides the three positions more or less fixed, what seems to be highly conserved is the overall succession of amino acids with specific biochemical properties that allows the correct folding of the domain and consequently the interaction with Groucho. In particular charged amino acids at position 4 and 5 (between the two Ile) and a serine/threonine in position 2 (following the conserved Phenylalanine residue) seems to be important for the stability of the short amphipathic α-helix in which the core of the domain is organized (for example through the formation of hydrogen bonds; Goldestein *et al.*, 2005; Jennings *et al.*, 2006). In addition to these structural properties, the aforementioned residues could also have a functional significance. For instance, the phosphorylation of Ser/Thr amino acid at position 2 could affect the Groucho binding capability, thus modulating the repression activity of the protein, as indicated by *in vitro* assays (Goldestein *et al.*, 2005).

Concerning specifically Gsx family proteins, it is already reported that also in this case the eh1 domain matches the generic consensus F-x-B-x-x-B, given that the first Phe residue is conserved and the two Ile are replaced by Val at position 3 and Leu at position 6. However, the other amino acids previously mentioned for their overall importance in the motif folding/Groucho binding (e.g. Ser/Thr phosphoacceptors and charged amino acids between the two B residues) seems all conserved (Von Ohlen *et al.*, 2007a). Furthermore, the repressive capability mediated by the eh1 domain, through Groucho interaction, has been demonstrated for ind, the ortholog of Gsx in *Drosophila* (Von Ohlen *et al.*, 2007a).

The comparative analysis that I performed through the alignment of several Gsx homologs, including the one from *Ciona intestinalis*, allowed me to verify that Ci-gsx protein sequence showed an high degree of conservation not only in the homeodomain (DNA-binding domain), and also in the N-terminal, at the level of Eh1-like motif. All the features that have been so far described for the "general" Eh1 peptides were indeed present, such as the conserved Phe residue, the phosphoacceptor serine at position 2, the negatively charged amino acid (glutamate, E) lying at position 4 as well as the generic B residues (positions 3 and 6). In addition, two serines in positions -1 and 5 were highly conserved in the Gsx sequences analyzed, including

Ci-gsx, but not in the other Eh1-containing transcription factors (my results compared with Goldestein *et al.*, 2005); analogously a basic residue, that is histidine (H) in *Ciona* and arginine (R) in other cases, was conserved at position -2 in Gsx deduced proteins.

On the basis of the similarity of this peptide sequence with both the "generic" Eh-1 motif and the Eh-1 motif found in the other Gsx protein from other species, it was possible to suppose that also in *Ciona intestinalis* Ci-gsx acts as a transcriptional repressor. This hypothesis was confirmed by the functional data presented in the paragraph 3.4, as later discussed.

The presence of the Eh-1 motif also suggests that *Ci-gsx* could have the potentiality to interact with Groucho to repress the target genes. In *Ciona* genome two Groucho homologs have been found (Satou *et al.*, 2003; Bajoghli, 2007), but besides some rough expression data in ANISEED database, no further information are available on these genes. Thus, as first step, we have already planned double in situ Ci-gsx/Groucho in order to delineate any overlapping expression territories that could be indicative of a potential interaction of these two factors during Ciona development.

4.2.2 Ci-gsx functional studies

In order to investigate the role of Ci-gsx in Ciona CNS, and in particular in photoreceptor cells differentiation, I used the lineage specific interference approach. To this end I exploited Ci-gsx its own promoter for a targeted expression of two modified Ci-gsx proteins, the constitutive activator (*GsxHD>VP16*) and repressor (*GsxHD>WRPW*) forms.

The overexpression of fusion constructs between the DNA-binding domain of a given transcription factor and the VP16 or WRPW sequences has been indeed proved to be very efficient in *Ciona intestinalis*, and very useful in order to unravel the function of the investigated factor (Davidson *et al.*, 2006; Beh *et al.*, 2007; Squarzoni *et al.*, 2012) A similar approach has also been applied in *Xenopus tropicalis* in the study of Gsh2 function (Winterbottom *et al.*, 2010).

The interference with photoreceptor differentiation usually do not result in any morphological alteration easily identifiable by microscopic observation. Thus, how to assess Ci-gsx phenotype? Here *Ci-Rx* revealed to be a key tool since 1) it is fundamental in the pathway of *Ciona* photoreceptor cells differentiation (D'Aniello *et al.*, 2006); 2) it is expressed later than *Ci-gsx* during embryogenesis (starting from early tailbud stage); 3) it is included in *Ci-gsx* expression domains. Thereby any altered *Ci-Rx* expression pattern, after targeted interference with Ci-gsx function, would undoubtedly indicate an involvement of Ci-gsx in the molecular pathway leading to ocellus photoreceptor differentiation.

This analysis has been conducted studying the effect of the interfering Ci-gsx constructs on the expression of *Ci-Rx* gene by both, whole mount *in situ* hybridization and qPCR experiments. Furthermore, it was tested also the capability of *Ci-Rx* promoter (*pRx*) to drive the expression of the reporter GFP in photoreceptor territories, when the *pRx>GFP* construct was co-electroporated with the chimeric Ci-gsx transgenes .

In all cases the results indicated a clear-cut downregulation of *Ci-Rx* expression and *pRx* activity following the expression of the constitutive transcriptional activator form of Ci-gsx (*GsxHD>VP16*).

The *in situ* data on the downregulation of *Ci-Rx* were very clear in both analyzed stages (middle and late tailbud stages) and strongly corroborated by the observation of *pRx* driven GFP fluorescence. Indeed, while almost 50% of control embryos were able to express GFP in the brain vesicle, though this regulatory region is non that strong, the percentage dropped down to 15% when *pRx>GFP* was co-electroporated with the constitutive active *pGsx>GsxHD>VP16* transgene. For what concerns the qPCR data, here too, the results were consistent with the previous observations.

In addition, in this experiment I tried to estimate also the variations in the expression level of *Ci-Arr*, a marker for differentiated photoreceptor cells (Nakagawa *et al.*, 2002) and *Ci-msxb*, a known Gsx target in *Drosophila* nervous system (Weiss *et al.*, 1998; Von Ohlen *et al.*, 2009).

Unfortunately I did not detect any significant change in the expression levels of both transcripts in the *pGsx>GsxHD>VP16* embryos compared to the controls.

Actually many technical issues have to be considered when analyzing these data. The first limiting factor is that the analysis was done on RNA extracted from the whole embryos that, at the late tailbud stage (when I collected them), are made of more than 1000 cells, with *Ci-Rx* being expressed, probably, in about 30 cells or less. On the other side *Ci-msxb* is instead expressed in many cells and in different tissue types, at the tailbud stage, so that a putative alteration in the little cellular population, theoretically regulated by Ci-gsx, is probably hidden. Finally, the tailbud stage could be too early to detect *Ci-Arr*, which at this stage is poorly expressed. In addition, even if the efficiency of the transgenesis driven by pGsx is really high (90% of the embryos), there is still a percentage of non-electroporated embryos that are included in the analysis. These not transgenic, wild type embryos, that cannot be sorted out versus the transgenic embryos, further contribute to dilute the mRNA used for the analysis. It is also important to note that the trend was steadily the same in more biological replicates. Therefore, based on all these remarks, it is reasonable to consider "significant" our fold change for *Ci-Rx* transcript, even if it is slightly less than 1,5.

As expected also from the protein sequence analysis, these results reasonably confirmed that also in *Ciona intestinalis* Ci-gsx acts as a transcriptional repressor, given the antimorphic effect of the GsxHD>VP16 construct. Furthermore, since this constitutive activator form of Ci-gsx led to a downregulation of *Ci-Rx*, the interaction between the two genes cannot be direct.

Reflecting on these findings, one could assume that the constitutive active form of Ci-gsx induces the transcription of other factor(s) which are able to repress (directly or indirectly) *Ci-Rx* transcription. Normally, Ci-gsx presence in Ci-Rx territories would block the transcription of these repressive factors, so that *Ci-Rx* can be expressed. The removal of Ci-gsx mediated block, through the VP16 domain, would conversely lead to the transcription of these unknown proteins. These factors, if present, could induce the downregulation of *Ci-Rx* and, consequentially, the block of photoreceptor cell differentiation. We cannot exclude also that the

presence of these factors could promote alternative fates of development in these presumed "photoreceptor" cell lineages. In this scenario *Ci-gsx* would not appear as a master gene, in the sense of a gene fundamental to direct a cell or group of cells to a specific fate, like for example *Otx* for CNS (Bertrand *et al.*, 2003), *Brachyury* for notochord (Corbo *et al.*, 1997), Ci-*ttf1* for endoderm (Ristoratore *et al.*, 1999) in *Ciona*. Rather *Ci-gsx* would behave as a "gregarius" but fundamental factor, that helps masters genes in directing the cells, where it is present, to a specific fate. This could be accomplished also by a further property of Gsx gene, that consists in controlling the balance between cell proliferation and differentiation, as already demonstrated in vertebrates. Hence, by exerting both activities, Ci-gsx would regulate the progression and the timing of photoreceptor cell differentiation in *Ciona*.

Thus the tendency was a downregulation of *Ci-Rx* transcript in *pGsx>GsxHD>VP16* electroporated embryos. It is noteworthy that an upregulation of *Ci-Rx* was instead detected in the opposite condition, in the *pGsx>GsxHD>WRPW* electroporated embryos. This data was somewhat in agreement also with the analysis of the pRx activity, since the fluorescent signal was observed in 90% of *pGsx>GsxHD>WRPW* transgenic larvae and appeared much stronger in the intensity, compared to the 50%, less intense, controls. In addition to these findings, it is remarkable that, at the late tailbud stage, the *pGsx>GsxHD>WRPW* embryos showed, in some case, a bilateral *Ci-Rx* expression, differently from the control embryos, in which *Ci-Rx* was kept, as usually, only on the right side of the sensory vesicle, where photoreceptor cells differentiate (***figure 3.11, B*** compared to ***A***)

If endogenous Ci-gsx behaves already as a repressor, how can be explained the stronger "power" on *Ci-Rx* activity, exerted by the "equally repressor" form GsxHD>WRPW? Firstly it is important to point out that these data are preliminary and needs to be confirmed. One can tentatively speculate that the exogenous *GsxHD>WRPW* mRNA, and thus the corresponding protein, is more persistent and much more copious with respect to the endogenous *Ci-gsx*. This could justify, somehow, the persistence of *Ci-Rx* signal on the left side and the higher percentage of fluorescent transgenic larvae. Another explanation could be related to the

different properties of the WRPW motif (which I artificially inserted in the protein) respect to the eh1 domain (which is normally present in Ci-gsx) as repressive moieties. Both of them are able to bind the co-repressor Groucho, but they are not simply interchangeable, equivalent sequences and their efficiency varies according to the protein structure and the biological context (Goldestein *et al.*, 2005; Buscarlet *et al.*, 2008; Jennings and Ish-Horowicz, 2008); in addition, the eh1 domain can undergo phosphorylation processes which regulate its binding to Groucho (Goldstein *et al.*, 2005). Taking all this into consideration, one can hypothesize that the WRPW might not interact with Groucho exactly as the wild type eh1 domain does, in the sense that the WRPW binding with Groucho could be stronger, or that the biological mechanisms which modulates the interaction WRPW/Groucho and the effects of this interaction are different compared to what normally occurs in the physiological regulation of Ci-gsx activity, for example on the left side of the embryo.

On these grounds, the overexpression of the wild type sequence of Ci-gsx under the control of its own promoter has been already scheduled in the lab.

In addition it would be really interesting to investigate the relation between Ci-gsx, Ci-Rx and Nodal signaling. Nodal pathway, indeed, plays a crucial role in the determination of the establishment of the left-right symmetry in various animals. Yoshida and Saiga (2011) recently demonstrated that Ci-Rx is negatively regulated by the Nodal signaling on the left side of the SV, when Ci-Nodal starts to be detectable in this region. Is there any relation between Ci-gsx and Ci-Nodal in controlling Ci-Rx disappearance? A preliminary approach to clarify Nodal/Ci-gsx relationship could consist in the analysis of *Ci-gsx* expression after interfering with Nodal function, 1) by using a Nodal signaling pathway inhibitor (SB431542, a strategy similar to the one applied in Yoshida and Saiga, 2001) and 2) by the overexpression of Nodal in the whole *Ciona* central nervous system through a promoter-guided strategy, using *Ci-Etr* enhancer (*pEtr*; D'Aniello *et al.*, 2011). In addition, it could be interesting to verify if the expression of *pGsx>GsxHD>VP16/pGsx>GsxHD>WRPW* is able to counteract the effect of Nodal inhibition/overexpression on *Ci-Rx* expression, reported in Yoshida and Saiga, 2001.

4.3 Analysis of *Ci-gsx* regulatory region

As previously mentioned, the isolation of *Ci-gsx* regulatory region represented the preliminary step of my PhD work, useful both for the analysis of Ci-gsx role in *Ciona* CNS development, as described above, and for the study of its transcriptional regulation. Studies on transcriptional regulation in *Ciona intestinalis* are facilitated by the compact genome of this organism, in which the *cis*-regulatory regions are usually in close proximity to the genes they control (Corbo *et al.*, 1997; Takahashi *et al.*, 1999; Fanelli *et al.* 2003; Alfano *et al.*, 2007; Squarzoni *et al.*, 2011).

In order to start the analysis of Ci-gsx transcriptional regulation, I firstly cloned a 5' genomic region of 2.8 kb upstream from the reporter gene mCherry (*pGsx>mCherry*), which resulted able to recapitulate the endogenous *Ci-gsx* pattern from late neurula to larval stages. The slight temporal lag between *Ci-gsx* mRNA appearance (gastrula stage) and the mCherry protein detection (late neurula) was because at the gastrula stage the amount of mCherry protein manufactured by this time was not enough to produce a visible signal under the fluorescence microscope. Actually, the transcribed mCherry reporter message was detected, by *in situ* hybridization, at the time of the appearance of the endogenous *Ci-gsx* transcript (data not shown). On the other side, the reporter signal persisted up to the larval stage, when the endogenous transcript is usually not visible through *in situ* experiments, thanks to the stability of mCherry protein compared with *Ci-gsx* transcript.

To narrow down the promoter, I took advantage of the comparison of the genomic sequences of *Ciona intestinalis* and its sister species, *Ciona savignyi.* The two species show indeed a high degree of conservation that allows to identify conserved regulatory sequences, given that non-coding DNA with regulatory functions has a higher tendency to be conserved with respect to other non-coding regions (Satoh *et al.*, 2003; Alfano *et al.*, 2007; Squarzoni *et al.*, 2011). So, prediction of cis-regulatory DNA sequences is facilitated by phylogenetic footprinting between these two genome sequences. The alignment of the 5' upstream region of *Ci-gsx* with the corresponding region of *C. savignyi* revealed three conserved fragments of about 200 bp, named I, II and III. Different combinations of conserved and non-conserved regions have been cloned

upstream from mCherry in order to attempt to determine the biological role of the three conserved regions.

In all the working fragments the signal remained localized in *Ci-gsx* territories, meaning that all of them recapitulated the endogenous expression, with higher or lower efficiency. This so detailed analysis of *Ci-gsx* promoter clearly indicated that all the elements, conserved and non-conserved, were required together to have the same percentage of expression (and with the same intensity) observed when electroporating the whole *pGsx* region.

As one could easily expect, considering not only the complexity and the dynamism of *Ci-gsx* expression pattern during the development, but also the strength and the efficiency of its promoter, it is clear that *Ci-gsx* transcriptional regulation cannot be related to a single small fragment. More probably, many elements are necessary for the fine regulation of Ci-gsx, contributing to the full activation of the gene according to the time and regions during embryogenesis.

On the other hand, the minimal information for *Ci-gsx* transcriptional activation was located in *pGsxIII* fragment, which seemed necessary and sufficient for the proper expression of the reporter gene in 70% (40%+++ plus 30%+) of the observed larvae. Accordingly *pGsxI* and p*GsxD* in which region III is lacking, were not active.

Based on these analyses, it is clear that the construct *pGsxIII* was the most suitable to carry this study forward.

4.3.1 Potential involvement of Sox transcription factors in Ci-gsx transcriptional activation

With a bioinformatic approach on *pGsxIII* region, I noticed three putative binding sites for member of the Sox family proteins and I named them sx0, sx1 and sx2; they had an organized pattern in their distribution, with a distance of 80/90 bp between them.

Sox transcription factors are found in all metazoan species and are characterized by the presence of a specific DNA binding domain, called high-mobility group (HMG) domain, highly conserved among Sox factors, that recognize a similar binding motif, 5'-(A/T)-(A/T)-C-A-A-(A/T)-G-3' (Wilson and Koopman, 2002). On the basis of the sequence similarity, ten groups

named SoxA to SoxH have been identified (Wegner, 1999; Schepers *et al.*, 2002; Wilson and Koopman, 2002). Sox proteins regulate countless developmental events ranging from gastrulation, haematopoietic and nervous system formation, development of skeleton, gonad, spleen, heart, blood vessels, melanocytes and so on (for reviews see: Wegner, 1999; Bowles *et al.*, 2000; Wegner, 2005; Kiefer, 2007). Accordingly, Sox genes are expressed in many tissues (Bowles *et al.*, 2000), and many genes must be under their transcriptional regulation, even if the properties of these proteins make this study quite difficult (Wegner, 2005). Sox transcription factors require indeed other protein partners for efficient target gene activation, forming multi-protein complexes, which are likely involved in determining cell specificity of Sox proteins (Wilson and Koopman, 2002; Wegner, 2005). Furthermore, Sox proteins are peculiar because they introduce a strong bend into DNA upon binding the minor grove, so that they act as architectural proteins altering the DNA three-dimensional conformation. In this way they may actually determine how other transcription factors interact with each other on the promoter of the target genes, and which transcriptional co-factors are recruited (Wegner, 2005).

Sox factors can act as transcriptional activator or repressors, according to the class to which they belong (Kiefer, 2007), but also depending upon their partners in a particular cell type (Wilson and Koopman, 2002).

In *Ciona intestinalis* seven Sox proteins have been identified, with representatives from the group B to F (Yamada *et al.*, 2003; Leveugle *et al.*, 2004). The expression pattern collected in ANISEED (http://www.aniseed.cnrs.fr/aniseed/) allowed me to select, among them, Ci-SoxB1 and Ci-SoxC as interesting candidates potentially involved in *Ci-gsx* expression. In addition, double *in situ* experiments that I performed to compare Ci-SoxB1 and Ci-gsx expression, confirmed that they co-localized since very early stages of development, making possible a regulative relationship between them.

Unfortunately, nothing is known about the role of Ci-SoxB1 and Ci-SoxC in ascidians CNS, besides the expression pattern of SoxB1 traced in *H. roretzi* (Miya and Nishida, 2003). In this organism *Hr-SoxB1* is specifically expressed in neural tissues during embryogenesis, in

particular in the neural plate in early stages and then, after neural tube closure, in the nerve cord and anterior neural tissues. This expression suggests a role in neural development, but this function has not been further investigated.

Concerning the role of SoxB1 and SoxC in other species, it results particularly interesting, especially in light of my hypotheses about Ci-gsx function.

SoxB1 proteins are indeed expressed in most neural precursors both in vertebrates and in *Drosophila* (for reviews about SoxB1 expression and functions see Kiefer, 2007; Guth and Wegner, 2008). In *Drosophila* SoxNeuro, belonging to the SoxB class, has a prominent role in neuroectodermal progenitors and neuroblasts formation. In vertebrates, the members of SoxB1 class (Sox1, Sox2 and Sox3) maintain neural cells in a pluripotent progenitor state counteracting neuronal differentiation, preventing cell cycle exit by repressing the activity of proneural genes. Later on in development, SoxB1 genes expression become restricted to particular CNS regions and the relative proteins acquire distinct functions upon differentiation, so that these transcription factors are required also for neuronal specification, maturation and terminal differentiation.

SoxC proteins (Sox4, Sox11 and Sox12) are less studied with respect to SoxB1, but it is known that they are present in post-mitotic differentiating neurons, regulating the latter steps of neuronal development. Furthermore, in mouse it has been demonstrated that SoxB1 and SoxC share a high number of target genes, meaning that different Sox proteins could sequentially act during CNS development, in order to coordinate neural lineage specific gene expression from early pluripotent stem cells to neuronal and glial gene expression (Bergsland *et al.*, 2011).

Taking all this into consideration, it seems that in general SoxB1 and SoxC genes could be framed into the molecular pathways that include Ci-gsx in *Ciona*,

To validate the importance of Sox binding sites predicted *in silico* on *pGsxIII*, I used two different approaches: a further 5' deletion of the fragment, thus generating the constructs *pGsx12>mCherry*, *pGsx13>mCherry* and *pGsx14>mCherry* (***figure 3.19***), and a site-directed mutagenesis.

The deletion study indicated that the 5' region of *pGsxIII* fragment, lacking in *pGsx12*, is probably very important for the activation of the reporter gene, since there was a strong reduction in the percentage of fluorescent (+++) larvae electroporated with *pGsx12>mCherry* (9%), compared to the ones electroporated with *pGsxIII>mCherry* (40%).

When considering *pGsx13>mCherry*, the promoter activity (6% of +++ fluorescent larvae) was less affected compared to *pGsx12>mCherry*, possibly suggesting that sx1 could be less important for the activation of the promoter. The shortest fragment, *pGsx14>mCherry*, was instead totally inactive.

Concerning the site-directed mutagenesis of the three putative Sox binding sites (***figure 3.20***), I observed that single mutation of each one of them resulted in the same reduction of promoter activity. Here too, the effect seemed barely stronger in Mut0_Sox, in agreement with the previous result.

The double mutant, in which both sx1 and sx2 sites were changed, showed a stronger decrease of activity, compared to the single mutations, but surprisingly the effect was not additive.

The triple mutant (*TrM_Sox*) deeply damaged the promoter activity, but it was still able to drive the expression of the reporter gene in 20% of the observed larvae, suggesting that other factors are probably also involved in this activation.

One explanation of these data could be related to the general properties of the Sox family genes. Assuming, indeed, that Sox factors are really involved in *Ci-gsx* activation, it is not surprising to imagine that the interaction with other proteins is necessary for the complete activation of the downstream gene, since, as previously mentioned, this mechanism is quite common for this class of transcription factors.

In addition, it has been already mentioned that often Sox binding induces an alteration of the three-dimensional DNA structure, which is necessary for the correct interaction with specific activators or repressors. In my experiments changes in the three-dimensional organization of the promoter region are not predictable nor verifiable. This fact constitutes a variable that could possibly justify the unexpected results obtained for example when electroporating the double

mutant. If the architecture of *pGsxIII* is affected in Sox mutants, it is possible that the single mutation interferes also with the binding of other factors in different regions of the promoter. Therefore, one could speculate that the effect of multiple mutations was not precisely additive because the alteration of three-dimensional structure, caused even by a single mutation, makes its effect stronger.

Considering all these elements, it seems that Sox proteins co-operate in order to activate *Ci-gsx* expression during *Ciona* development and multiple binding on the promoter is required for its complete transcriptional activation. This result is not unexpected, given that the co-operation of multiple binding sites for the activation of a given transcription factor has been already described in *Ciona* (Bertrand *et al.*, 2003; Squarzoni *et al.*, 2011).

In addition, at least another factor is likely involved in this process.

Of course these are just hypotheses and the results I have obtained so far are preliminary indications and need to be corroborated by the experimental proof of the direct interaction between one of the Sox proteins and *Ci-gsx* promoter.

In order to study the direct interaction between *Ci-gsx* promoter and Ci-SoxB1 or Ci-SoxC transcription factors, an Electrophoretic Mobility Shift Assay could be required (EMSA). Regrettably, at the moment, Stazione Zoologica has no permission to use the radioactive isotopes required for this experiment.

An alternative strategy, which has been successfully applied by our research group (D'Aniello *et al.*, 2011), could consist in testing the capability of Sox proteins, ectopically expressed in the notochord by using *Ci-Brachyury* promoter (Corbo *et al.*, 1997), to activate the expression of a reporter gene (mCherry o GFP), under the control of pGsxIII promoter, just in the notochord. The detection of *pGsxIII* promoter activity would indicate that Sox regulates Ci-gsx. Nonetheless, since our data indicate that Sox family genes are probably not the sole factors sufficient for Ci-gsx transcriptional activation, this experiment is unlikely to work if the other factor(s) cooperating with Sox are missed.

Comparing the results obtained by the electroporation of *TrM_Sox>mCherry* and *pGsx12DM>mCherry*, it was quite evident that the region included between sx1 and sx2 binding sites (corresponding to the region included between *pGsx13* and *pGsx14*) contains fundamental information for the full *Ci-gsx* transcriptional activation.

In order to find the factor(s) probably interacting with Sox, I decided to further restrict *Ci-gsx* regulatory region. I identified a fragment of 185 bp, (construct *pGsx13a2>mCherry*), which is still able to drive the expression of the reporter gene in 3% (+++) 22% (+) of the observed larvae. The analysis of this region, compared with the one immediately shorter (181 bp) contained in the plasmid *pGsx13b>mCherry*, will be probably useful in the identification of the missing factor. The construct *pGsx13b>mCherry*, although containing the last putative sox binding site, sx2, was not able to induce mCherry expression. It is therefore possible that *pGsx13a2* contains the binding site for "the missed factor(s)". A first rough bioinformatic scan on this region, using both JASPAR and Genomatix databases, did not allow me to find any interesting candidate and a more detailed study is of course required. Unfortunately, I had to momentarily suspend this analysis because of limitation of time. The site-directed mutagenesis of the four bases included between *pGsx13a2* and *pGsx13b* regions will probably confirm the importance of this region for *Ci-gsx* transcriptional activation.

Interestingly, no one of the fragments tested so far leads to an ectopic expression of the reporter gene. Furthermore, as shown in ***figure 3.17*** and ***3.18***, the expression of *Ci-SoxB1* and *Ci-SoxC* in *Ciona* nervous system is wider than the one observed for *Ci-gsx*. On these bases, it is possible to suppose that the correct *Ci-gsx* expression could be due to a transcriptional co-activator expressed only in *Ci-gsx* territories and/or to a transcriptional repressor able to downregulate the gene outside its wild type domain. The binding site for this putative repressor seems present in *pGsx13a2* region, which is still able to drive the correct mCherry expression.

4.3.2 Potential involvement of Ci-Msxb in Ci-gsx transcriptional repression

Together with the three putative Sox binding sites, the bioinformatic analysis of *pGsxIII* also revealed the presence of two potential recognition sites for Msx.

This results was intriguing, given that in *Drosophila* a cross repressive interaction between msh (Msx homolog) and Ind has been described (Weiss *et al.*, 1998; Von Ohlen *et al.*, 2009; Seibert and Urbach, 2010). It is indeed known that *msh* is a direct target of Ind in the developing trunk neuroectoderm, in a hierarchical cascade of transcriptional repression, necessary to establish and maintain the sharp boundaries between their domains of expression (Cowden and Levine, 2003). However, the expression of these genes exhibits segment specific differences in the procephalic neuroectoderm and in brain neuroblasts, where the repressive interaction between msh and ind is reciprocal, so ind is also a target of msh (Seibert and Urbach, 2010).

This direct relationship between Msx and Gsx was not confirmed in vertebrates, and in particular in *Xenopus tropicalis*, where Winterbottom *et al.* (2010), investigated this aspect. In the frog Gsh2 and Msx1 are expressed in the neural plate in similar medio-lateral position to that of the Drosophila genes *ind* and *msh* (Illes *et al.*, 2009). Nevertheless, Gsh2 overexpression in the whole embryo causes a downregulation of *Msx1* anteriorly in the neural plate, and a general expansion of *Msx1* domains in more posterior regions. In addition the antimorphic Gsh2-VP16 downregulates, rather than upregulates, *Msx1*, thus arguing against a direct interaction. It seems therefore that, during vertebrate evolution, the components of the ancestral neural patterning system are generally conserved, but the interactions between these transcription factors can vary even within a single organism between different body regions and over the course of development (Winterbottom *et al.*, 2010). In order to add information to this scenario, the analysis of these interactions in ascidians, which are closely related to vertebrates, will be really useful.

In *Ciona intestinalis* one *Msx* ortholog has been found, *Ci-msxb*, characterized by a complex expression pattern during embryogenesis, including, at the beginning of gastrulation, the precursors of mesenchyme cells, muscle, spinal cord, endodermal strand, nervous system,

pigment cells and primordial pharynx. In subsequent stages of development the transcript is restricted in smaller area, including the sensory organs, the neck and the primordial pharynx (Aniello *et al.*, 1999).

Double *in situ* hybridization experiments demonstrated that the *Ci-gsx* and *Ci-msxb* occupy adjacent territories in early developmental stages, similarly to what has been observed in *Drosophila*. In particular the cells in which *Ci-gsx* was expressed (in particular at gastrula and neurula stages) were located next to pigment cells precursors, where *Ci-msxb* was expressed. This could suggest a putative interaction between these two factors and encouraged me to proceed in this direction.

The ectopic expression of Ci-msxb protein in *Ci-gsx* territories, driven by *pGsx* promoter, showed indeed a clear reduction both of endogenous *Ci-gsx* expression, as revealed by WMISH assays, and of *pGsx* activity, as demonstrated by co-electroporation experiments of *pGsx>Msx* together with *pGsxIII>mCherry* (***figure 3.24***). This result suggests that Ci-msxb could downregulate *Ci-gsx*. I also prepared different constructs in which Ci-msxb coding sequence was under the control of *Ci-Etr* enhancer (*pEtr*), active throughout *Ciona* central nervous system, including *Ci-gsx* territories, and already successfully used in our group for similar purposes (D'Aniello *et al.*, 2011). These constructs will be soon tested.

These preliminary data support the idea of a possible Ci-msxb role in the repression of Ci-gsx, maybe through the binding at least to one of the two Msx sites predicted *in silico*, and here named *msx1* and *msx2*. Interestingly, the constructs in which *msx1* putative binding site is absent (***figure 3.21***, compared with ***figure 3.22***) did not lead to an ectopic expression of *mCherry* reporter gene, whereas *msx2* is always present in the fragments that I tested so far. Therefore, it is possible to predict that *msx2* putative binding site could be more important than *msx1* for *Ci-gsx* repression. This suppression could act by downregulating *Ci-gsx* in pigment cells precursors, a mechanism that could be perhaps required to maintain their correct identity. Ongoing experiments in the lab are just devoted to confirm the importance of the two putative binding sites for Ci-msxb on *pGsxIII* by a site-directed mutagenesis.

4.4 Conclusions and future directions

In sum, the results I am reporting clearly indicate that *Ci-gsx*, whose expression domain includes the ocellus photoreceptor cells precursors since gastrula stage, plays a role in the molecular pathway controlling their differentiation, acting as a transcriptional repressor. Targeted perturbation approaches revealed indeed that Ci-gsx activity is required for the correct expression of *Ci-Rx*, a gene fundamental for ocellus development (D'Aniello *et al.*, 2006).

These data that appear to be innovative and interesting in the literary landscape. It is indeed the first time that a ParaHox gene has been linked to the "visual" system development, moreover in an organism closely related to vertebrates. In addition, I started to shed light on the early steps of ocellus photoreceptor cells specification and differentiation, a process still largely unknown in ascidian community, despite its relevance from a developmental and evolutionary point of view. I also demonstrated that *Ci-gsx* expression is not limited to the ocellus photoreceptor cells precursors, thus allowing to suppose that its function is required for the formation of several cell types in the presumptive sensory vesicle (examples could be the dopaminergic coronet cells and/or the group III of photoreceptor cells).

I provided a detailed analysis of the 5' *Ci-gsx* regulatory region, aimed at the characterization of the specific module(s) responsible for *Ci-gsx* activation. As one could easily expect from the complexity and the dynamism of *Ci-gsx* expression pattern during *Ciona* development, its fine transcriptional regulation turned out to be difficult to be extricated. However, my studies permitted to identify specific Sox modules in *pGsx* promoter which are involved in *Ci-gsx* transcriptional activation. Thus a member of Sox family is likely to be a *trans*-acting protein, that would exert its action most probably in combination with other factor(s), in order to drive a full and complete *Ci-gsx* expression in the right territories and at the right developmental time. To confirm this preliminary data will permit to identify for the first time a direct Sox target in *Ciona intestinalis* central nervous system development.

My data also indicated that in *Ciona* CNS a repressive mechanisms, defining the correct identity of adjacent territories, ideally similar to the one observed in *Drosophila* and vertebrates, could act to block *Ci-gsx* expression in pigment cell precursors. I provided indeed interesting

evidences of a repressive action of Ci-msxb, a gene which is expressed, among different territories, also in pigment cell precursors, on *Ci-gsx*, thus opening a new scenario about the relationship between the components of the ancestral neural patterning system.

In order to confirm the results obtained so far, but mainly in order to have a wider picture of the molecular mechanisms in which *Ci-gsx* is involved, a transcriptomic approach, using a microarray analysis (or alternatively deep sequencing) has been already planned.

To this aim, the preliminary and essential step will be the FACS (Fluorescence Activated Cell Sorting), through which it is possible to select target cells from the whole embryo, on the basis of transgenic fluorescence. Briefly, transgenic embryos co-electroporated with *pGsx>GFP* + *pGsx>GsxHD>VP16* or *pGsx>GFP* + *pGsx>GsxHD>WRPW*, in parallel with control embryos, electroporated with *pGsx>GFP*, will be dissociated to obtain a cell suspension for the FACS. With this technology, the GFP positive population, constituted by the cells of *Ci-gsx* lineage that are expressing the exogenous DNA, will be "fished" from the remaining GFP negative ones. Total RNA extracted from these cells will permit a lineage specific transcription profiling, using microarray or deep sequencing, as previously mentioned. From the analysis of genes that are upregulated or downregulated in the two different conditions with respect to the control, it will be hopefully possible to find Ci-gsx targets, in order to further clarify the role of this transcription factor in *Ciona* nervous system development. Moreover, this study will hopefully permit to reveal the gene regulatory network in which Ci-gsx is involved, thus supplying new insights regarding the role, still controversial, of *Gsx* gene in the ancestor of bilaterians.

From this perspective it is also important to investigate the regulatory control of Gsx expression in different species, in order to reveal whether there are or not comparable regulatory networks that would be consistent with conserved, complicated or simplified expression of Gsx in each lineage. The number of experiments already scheduled and previously indicated, devoted to the identification of the transcriptional complex allowing *Ci-gsx* fine regulation, will be conducted in parallel with the functional/transcriptomic approach just described.

The continuation of my PhD work will therefore permit to have valuable information, in a model system closely related to vertebrates, about the function and the mechanism controlling the expression of *Ci-gsx*, a factor whose fundamental role in the developing nervous system of several organisms (including mammals) is becoming more and more interesting and intriguing.

REFERENCES

Alfano C, Teresa Russo M, Spagnuolo A. (2007). Developmental expression and transcriptional regulation of *Ci-Pans*, a novel neural marker gene of the ascidian *Ciona intestinalis*. *Gene*, 406: 36-41.

Aniello F, Locascio A, Villani MG, Di Gregorio A, Fucci L, Branno M. (1999). Identification and developmental expression of *Ci-msxb*: a novel homologue of *Drosophila msh* gene in *Ciona intestinalis*. *Mech Dev*, 88: 123-26.

Annunziata R, Martinez P, Arnone MI. (2013). Intact cluster and chordate-like expression of ParaHox genes in a sea star. *BMC Biol*, 11: 68.

Arnone MI, Rizzo F, Annunziata R, Cameron RA, Peterson KJ, Martinez P. (2006). Genetic organization and embryonic expression of the ParaHox genes in the sea urchin *S. purpuratus*: insights into the relationship between clustering and colinearity. *Dev Biol*, 300: 63-73.

Arshavsky VY. (2002). Rhodopsin phosphorylation: from terminating single photon responses to photoreceptor dark adaptation. *Trends Neurosci*. 25: 124-6.

Bailey TJ, El-Hodiri H, Zhang L, Shah R, Mathers PH, Jamrich M. (2004). Regulation of vertebrate eye development by *Rx* genes. *Int J Dev Biol*, 48: 761-70.

Bajoghli B. (2007). Evolution of the Groucho/Tle gene family: gene organization and duplication events. *Dev Genes Evol*, 217: 613-8.

Beaster-Jones L, Kaltenbach SL, Koop D, Yuan S, Chastain R, Holland LZ. (2008). Expression of somite segmentation genes in amphioxus: a clock without a wavefront? *Dev Genes Evol*, 218: 599-611.

Beh J, Shi W, Levine M, Davidson B, Christiaen L. (2007). FoxF is essential for FGF-induced migration of heart progenitor cells in the ascidian Ciona intestinalis. *Development*, 134: 3297-305.

Bergsland M, Ramsköld D, Zaouter C, Klum S, Sandberg R, Muhr J. (2011). Sequentially acting Sox transcription factors in neural lineage development. *Genes Dev*, 25: 2453-64.

Bertrand V, Hudson C, Caillol D, Popovici C, Lemaire P. (2003). Neural tissue in ascidian embryos is induced by FGF9/16/20, acting via a combination of maternal GATA and Ets transcription factors. *Cell*, 115: 615-27.

Blomhoff R. and Blomhoff HK. (2006). Overview of retinoid metabolism and function. *J Neurobiol*, 66: 606-30.

Bowles J, Schepers G, Koopman P. (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol*, 227: 239-55.

Brooke NM, Garcia-Fernandez J, Holland PW. (1998). The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature*, 392: 920-2.

Callaerts P, Halder G. Gehring WJ. (1997). PAX-6 in development and evolution. *Annu Rev Neurosci*, 20: 483-532.

Caracciolo A, Gesualdo I, Branno M, Aniello F, Di Lauro R, Palumbo A. (1997). Specific cellular localization of tyrosinase mRNA during *Ciona intestinalis* larval development. *Dev Growth Differ*, 39: 437-44.

Carney RS, Cocas LA, Hirata T, Mansfield K, Corbin JG. (2009). Differential regulation of telencephalic pallial-subpallial boundary patterning by Pax6 and Gsh2. Cereb Cortex, 19: 745-59.

Cheesman SE and Eisen JS. (2004). gsh1 demarcates hypothalamus and intermediate spinal cord in zebrafish. *Gene Expr Patterns*, 5:107-12.

Christiaen L, Wagner E, Shi W, Levine M. (2009). Whole-mount *in situ* hybridization on sea squirt (*Ciona intestinalis*) embryos. *Cold Spring Harb Protocol*, 12: pdb.prot5348.

Cole AG and Meinertzhagen IA (2004). The central nervous system of the ascidian larva: mitotic history of cells forming the neural tube in late embryonic *Ciona intestinalis*. *Dev Biol*, 271: 239-62.

Conklin EG (1905a). Mosaic development in ascidian egg. *J Exp Zool*, 2: 145-223.

Conklin EG. (1905b). Organ-forming substances in the eggs of ascidians. *Biol Bull*, 8: 205-30.

Conklin EG. (1905c). The organization and cell-lineage of the ascidian egg. *J Acad Nat Science*, 13: 1-119.

Corbo JC, Levine M, Zeller RW. (1997). Characterization of a notochord specific enhancer from the Brachyury promoter region of the ascidian, *Ciona intestinalis*. *Development*, *124*: 589-602.

Cowden J and Levine M. (2003). Ventral dominance governs sequential patterns of gene expression across the dorsal-ventral axis of the neuroectoderm in the *Drosophila* embryo. *Dev Biol*, 262: 335-49.

D'Aniello E, Pezzotti MR, Locascio A, Branno M. (2011). Onecut is a direct neural-specific transcriptional activator of Rx in *Ciona intestinalis*. *Dev Biol*, 355:358-71.

D'Aniello S, D'Aniello E, Locascio A, Memoli A, Corrado M, Russo MT, Aniello F, Fucci L, Brown ER, Branno M. (2006). The ascidian homolog of the vertebrate homeobox gene *Rx* is essential for ocellus development and function. *Differentiation*, 74: 222-34.

Dalafave DS. (2009). Prediction of functional engrailed homology-1 protein motif from sequence. *Bioinformation*, 4: 229-32.

Davidson B, Shi W, Beh J, Christiaen L, Levine M. (2006). FGF signaling delineates the cardiac progenitor field in the simple chordate, *Ciona intestinalis*. *Genes Dev*, 20: 2728-38.

Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A, et al., (2002). The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science*, 298: 2157-67.

del Marmol V and Beermann F. (1996). Tyrosinase and related proteins in mammalian pigmentation. *FEBS Lett*, 381: 165-8.

Delsuc F, Brinkmann H, Chourrout D, Philippe H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature*, 439: 965-8.

Denes AS, Jekely G, Steinmetz PR, Raible F, Snyman H, Prud'homme B. Ferrier DE, Balavoine G, Arendt D. (2007). Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in bilateria. *Cell*, 129: 277-88.

Deschet K, Bourrat F, Chourrout D, Joly JS. (1998). Expression domains of the medaka (*Oryzias latipes*) *Ol-Gsh 1* gene are reminiscent of those of clustered and orphan homeobox genes. *Dev Genes Evol*, 208: 235-44.

Dilly PN. (1961). Electron microscope observations of the receptors in the sensory vesicle of the ascidian tadpole. *Nature*, 191: 786-7.

Dilly PN. (1969a). The ultrastructure of the test of the tadpole larva of Ciona intestinalis. *Z Zellforsch Mikrosk Anat*, 95: 331-46.

Dilly PN. (1969b). Studies on the receptors in Ciona intestinalis. 3. A second type of photoreceptor in the tadpole larva of Ciona intestinalis. *Z Zellforsch Mikrosk Anat*, 96: 63-5.

Dufour H, Chettouh Z, Deyts C, de Rosa R, Goridis C, Joly JS, Brunet JF. (2006). Precraniate origin of cranial motoneurons. *Proc Natl Acad Sci USA*, 103: 8727-32.

Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD, Sorensen MV, Haddock SH, Schmidt-Rhaesa A, Okusu A, Kristensen RM, Wheeler WC, Martindale MQ, Giribet G. (2008). Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature*, 452: 745-9.

Eakin RM and Kuda A. (1971). Ultrastructure of sensory receptors in Ascidian tadpoles. *Z Zellforsch Mikrosk Anat*, 112: 287-312.

Esposito R, D'Aniello S, Squarzoni P, Pezzotti MR, Ristoratore F, Spagnuolo A. (2012). New insights into the evolution of metazoan tyrosinase gene family. *PLoS ONE*, 7(4): e35731.

Fanelli A, Lania G, Spagnuolo A, Di Lauro R. (2003). Interplay of negative and positive signals controls endoderm-specific expression of the ascidian *Cititf1* gene promoter. *Dev Biol*, 263: 12-23.

Ferrier DE and Holland PW. (2002). *Ciona intestinalis* ParaHox genes: evolution of Hox/ParaHox cluster integrity, developmental mode, and temporal colinearity. *Mol Phylogenet Evol*, 24: 412-7.

Ferrier DE, Dewar K, Cook A, Chang JL, Hill-Force A, Amemiya C. (2005). The chordate ParaHox cluster. *Curr Biol*, 15: R820–2.

Finnerty JR, Paulson D, Burton P, Pang K, Martindale MQ. (2003). Early evolution of a homeobox gene: the parahox gene Gsx in the Cnidaria and the Bilateria. Evol Dev, 5: 331-45.

Fleige S, Walf V, Huch S, Prgomet C, Sehm J, Pfaffl MW. (2006). Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR. *Biotechnol Lett*, 28: 1601-13.

Fröbius AC and Seaver EC. (2006). ParaHox gene expression in the polychaetes annelid *Capitella sp I*. *Dev Genes Evol*, 216: 81-8.

Furlong RF and Mulley JF. (2008). ParaHox cluster evolution–hagfish and beyond. *Zoolog Sci*, 25: 955-9.

Galliot B and Quiquand M. (2011). A two-step process in the emergence of neurogenesis. Eur J Neurosci, 34: 847-62.

Garstang M and Ferrier DEK. (2013). Time is of the essence for ParaHox homeobox gene clustering. *BMC Biol*, 11:72.

Gehring WJ. (2004). Historical perspective on the development and evolution of eyes and photoreceptors. *Int J Dev Biol*, 48:707-17.
Gehring WJ. (2005). New perspectives on eye development and the evolution of eyes and photoreceptors. *J Hered*, 96: 171-84.

Gehring WJ and Ikeo K. (1999). Pax6: mastering eye morphogenesis and eye evolution. *Trends Genet*, 15: 371-7.

Gionti M, Ristoratore F, Di Gregorio A, Aniello F, Branno M, Di Lauro R. (1998). *Cihox5*, a new *Ciona intestinalis* Hox-related gene, is involved in regionalization of the spinal cord. *Dev Genes Evol*, 207: 515-23.

Glardon S, Callaerts P, Halder G, Gehring WJ. (1997). Conservation of Pax-6 in a lower chordate, the ascidian *Phallusia mammillata*. *Development*, 124: 817-25.

Goldstein RE, Cook O, Dinur T, Pisante A, Karandikar UC, Bidwai A, Paroush Z. (2005). An eh1-like motif in odd-skipped mediates recruitment of Groucho and repression in vivo. *Mol Cell Biol*, 25: 10711-20.

Gorman AL, McReynolds JS, Barnes SN. (1971). Photoreceptors in primitive chordates: fine structure, hyperpolarizing receptor potentials, and evolution. *Science*, 172: 1052-4.

Guth SIE and Wegner M. (2008). Having it both ways: Sox protein function between conservation and innovation. *Cell Mol Life Sci*, 65: 3000-18.

Halder G, Callaerts P, Gehring WJ. (1995). Induction of ectopic eyes by targeted expression of the eyeless gene in *Drosophila*. *Science*, 267: 1788-92.

Holland LZ, Laudet V, Schubert M. (2004). The chordate amphioxus: an emerging model organism for developmental biology. *Cell Mol Life Sci*, 61: 2290-308.

Hong JW, Hendrix DA, Papatsenko D, Levine MS. (2008). How the Dorsal gradient works: insights from postgenome technologies. *Proc Natl Acad Sci USA*, 105: 20072-6.

Horie T, Orii H, Nakagawa M. (2005). Structure of ocellus photoreceptors in the ascidian *Ciona intestinalis* larva as revealed by an anti-arrestin antibody. *J Neurobiol*, 65: 241-50.

Horie T, Sakurai D, Ohtsuki H, Terakita A, Shichida Y, Usukura J, Kusakabe T, Tsuda M. (2008). Pigmented and nonpigmented ocelli in the brain vesicle of the ascidian larva. *J Comp Neurol*, 509: 88-102.

Hsieh-Li HM, Witte DP, Szucsik JC, Weinstein M, Li H, Potter SS. (1995). *Gsh-2*, a murine homeobox gene expressed in the developing brain. *Mech Dev*, 50: 177-86.

Hudson C and Lemaire P. (2001). Induction of anterior neural fates in the ascidian *Ciona intestinalis*. *Mech Dev*, 100:189-203.

Hudson C and Yasuo H. (2005). Patterning across the ascidian neural plate by lateral Nodal signalling sources. *Development*, 132: 1199-210.

Hudson C and Yasuo H. (2008). Similarity and diversity in mechanisms of muscle fate induction between ascidian species. *Biol Cell*, 100: 265-77.

Hudson C, Lotito S, Yasuo H. (2007). Sequential and combinatorial inputs from Nodal, Delta2/Notch and FGF/MEK/ERK signaling pathways establish a grid-like organisation of distinct cell identities in the ascidian neural plate. *Development*, 134: 3527-37.

Hui JH, Raible F, Korchagina N, Dray N, Samain S, Magdelenat G, Jubin C, Segurens B, Balavoine G, Arendt D, Ferrier DE. (2009). Features of the ancestral bilaterian inferred from *Platynereis dumerilii* ParaHox genes. *BMC Biol*, 7: 43.

Ikuta T, Chen Y-C, Annunziata R, Ting H-C, Tung C-h, Koyanagi R, Tagawa K, Humphreys T, Fujiyama A, Saiga H, Satoh N, Yu J-K, Arnone MI, Su Y-H. (2013). Identification of an intact ParaHox cluster with temporal colinearity but residual spatial colinearity in the hemichordate *Ptychodera flava*. *BMC Evol Biol*, 13:129.

Illes JC, Winterbottom E, Isaacs, HV. (2009). Cloning and expressionvanalysis of the anterior parahox genes, Gsh1 and Gsh2 from Xenopus tropicalis. *Dev Dyn*, 238: 194-203.

Imai JH and Meinertzhagen IA. (2007). Neurons of the ascidian larval nervous system in Ciona intestinalis: I. Central nervous system. *J Comp Neurol*, 501: 316-34.

Imai KS, Hino K, Yagi K, Satoh N, Satou Y. (2004). Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. *Development*, 131: 4047-58.

Imai KS, Levine M, Satoh N, Satou Y. (2006). Regulatory blueprint for a chordate embryo. Science, 312: 1183-7.

Imai KS, Satoh N, Satou Y. (2002). Region specific gene expressions in the central nervous system of the ascidian embryo. *Gene Expr Patterns*, 2: 319-21.

Imai KS, Stolfi A, Levine M, Satou Y. (2009). Gene regulatory networks underlying the compartmentalization of the *Ciona* central nervous system. *Development, 136*: 285-93.

Imbeaud S, Graudens E, Boulanger V, Barlet X, Zaborski P, Eveno E, Mueller O, Schroeder A, Auffray C. (2005). Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res*, 33:e56.

Jacobsen H, Klenow H, Overgaard-Hansen K. (1974). The N-terminal amino-acid sequences of DNA polymerase I from Escherichia coli and of the large and the small fragments obtained by a limited proteolysis. *Eur J Biochem*, 45: 623-7.

Jennings BH and Ish-Horowicz D. (2008). The Groucho/TLE/Grg family of transcriptional co-repressors. *Genome Biol*, 9: 205.

Jennings BH, Pickles LM, Wainwright SM, Roe SM, Pearl LH, Ish-Horowicz D. (2006). Molecular Recognition of Transcriptional Repressor Motifs by the WD Domain of the Groucho/TLE Corepressor. *Mol Cell*, 22: 645-55.

Jiménez G, Guichet A, Ephrussi A, Casanova J. (2000). Relief of gene repression by torso RTK signaling: Role of capicua in Drosophila terminal and dorsoventral patterning. *Genes Dev*, 14: 224-31.

Jiménez G, Paroush Z, Ish-Horowicz D. (1997). Groucho acts as a corepressor for a subset of negative regulators, including Hairy and Engrailed. *Genes Dev*, 11: 3072-82.

Jiménez G, Verrijzer CP, Ish-Horowicz D. (1999). A conserved motif in Goosecoid mediates Groucho-dependent repression in *Drosophila* embryos. *Mol Cell Biol*, 19: 2080-7.

Johnson DS, Davidson B, Brown CD, Smith WC, Sidow A. (2004). Noncoding regulatory sequences of Ciona exhibit strong correspondence between evolutionary constraint and functional importance. *Genome Res*, 14: 2448-56.

Katz MJ. (1983). Comparative anatomy of the tunicate tadpole, *Ciona intestinalis*. *Biol Bull*, 164: 1-27.

Kiefer J. (2007). Back to Basics: *Sox* Genes. *Dev Dyn*, 236: 2356-66.

Kulakova MA, Cook CE, Andreeva TF. (2008). ParaHox gene expression in larval and postlarval development of the polychaete *Nereis virens* (*Annelida, Lophotrochozoa*). *BMC Dev Biol*, 8: 61.

Kusakabe T, Kusakabe R, Kawakami I, Satou Y, Satoh N, Tsuda M. (2001). *Ci-opsin1*, a vertebrate-type opsin gene, expressed in the larval ocellus of the ascidian *Ciona intestinalis*. *FEBS Lett*, 506: 69-72.

Lamb TD, Collin SP, Pugh EN Jr. (2007). Evolution of the vertebrate eye: opsins, photoreceptors retina and eye cup. *Nat Rev Neurosci*, 8: 960-76.

Lemaire P. (2008). Unfolding a chordate developmental program, one cell at a time: Invariant cell lineages, short-range inductions and evolutionary plasticity in ascidians. *Dev Biol*, 332: 48-60.

Leong K, Brunet L, Berk AJ. (1988). Factors responsible for the higher transcriptional activity of extracts of adenovirus-infected cells fractionate with the TATA box transcription factor. *Mol. Cell. Biol*, 8: 1765-74.

Leveugle M, Prat K, Popovici C, Birnbaum D, Coulier F. (2004). Phylogenetic analysis of *Ciona intestinalis* gene superfamilies supports thehypothesis of successive gene expansions. *J Mol Evol*, 58: 168-81.

Lewin B. (2000). Homeodomains bind related targets in DNA. *In Genes VII. Oxford: Oxford University Pres*, 660-2.

Lim B, Samper N, Lu H, Rushlow C, Jiménez G, Shvartsamn SY. (2013.) Kinetics of gene derepression by ERK signaling. *Proc Natl Acad Sci USA*, 110: 10330-5.

Liu A and Joyner AL. (2001). Early anterior/posterior patterning of the midbrain and cerebellum. *Annu Rev Neurosci*, 24: 869-96.

Locascio A, Ristoratore F, Spagnuolo A, Zanetti L, Branno M. (2009). Genetic perspectives on the ascidian central nervous system. *Invertebrate Surviv J*, 6: S35-S45.

MacDonald PM, Ingham P, Struhl G. (1986). Isolation, structure, and expression of *even-skipped*: A second pair-rule gene of *Drosophila* containing a homeobox. *Cell*, 47: 721–34.

Marikawa Y, Yoshida S, Satoh, N. (1994). Development of egg fragments of the ascidian *Ciona* savignyi: the cytoplasmic factors responsible for muscle differentiation are separated into a specific fragment. *Dev biol*, 162: 134-42.

Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, Pachter LS, Dubchak I. (2000). VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics*, 16: 1046-7

Mazet F, Hutt JA, Millard J, Shimeld SM. (2003). *Pax* gene expression in the developing central nervous system of *Ciona intestinalis*. *Gene Expr Patterns*, 3: 743-5.

Mc Donald JA, Holbrook S, Isshiki T, Weiss J, Doe CQ, Mellerick DM. (1998). Dorso-ventral patterning in the *Drosophila* CNS: the *vnd* homeobox gene specifies ventral column identity. *Genes Dev*, 12: 3606-12.

McGinnis W, Levine MS, Hafen E, Kuroiwa A, Gehring WJ. (1984). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature*, 308: 428-33

Meinertzhagen IA, Okamura Y. (2001). The larval ascidian nervous system: the chordate brain from its small beginnings. *Trends Neurosci*, 24: 401-10.

Méndez-Gómez HR and Vicario-Abejón C. (2012). The homeobox gene Gsx2 regulates the self-renewal and differentiation of neural stem cells and the cell fate of postnatal progenitors. *PLoS One*, 7: e29799.

Mendivil Ramos O, Barker D, Ferrier DEK. (2012). Ghost loci imply Hox and ParaHox existence in the last common ancestor of animals. *Curr Biol*, 22: 1951-6.

Miljkovic-Licina M, Chera S, Ghila L, Galliot B. (2007). Head regeneration in wild-type hydra requires de novo neurogenesis. *Development*, 134: 1191-201.

Miya T and Nishida H. (2003). Expression pattern and transcriptional control of *SoxB1* in embryos of the ascidian *Halocynthia roretzi*. *Zool Sci*, 20: 59-67.

Moret F, Christiaen L, Deyts C, Blin M, Joly JS, Vernier P. (2005a). The dopamine-synthesizing cells in the swimming larva of the tunicate *Ciona intestinalis* are located only in the hypothalamus-related domain of the sensory vesicle. *Eur J Neurosci*, 21: 3043-55.

Moret F, Christiaen L, Deyts C, Blin M, Vernier P, Joly JS. (2005b). Regulatory gene expressions in the ascidian ventral sensory vesicle: evolutionary relationships with the vertebrate hypothalamus. *Dev biol*, 277: 567-79.

Mulley JF, Chiu CH, Holland PW. (2006). Breakup of a homeobox cluster after genome duplication in teleosts. *Proc Natl Acad Sci USA*, 103: 10369-72.

Munro E, Robin F, Lemaire P. (2006). Cellular morphogenesis in ascidians: how to shape a simple tadpole. *Curr Opin Genet Dev*, 16: 399-405.

Nakagawa M, Orii H, Yoshida N, Jojima E, Horie T, Yoshida R, Haga T, Tsuda M. (2002). Ascidian arrestin (*Ci-arr*), the origin of the visual and nonvisual arrestins of vertebrate. *Eur J Biochem*, 269: 5112-8.

Nakashima K, Yamada L, Satou Y, Azuma J and Satoh N. (2004). The evolutionary origin of animal cellulose synthase. *Dev Genes Evol*, 214: 81-8.

Nakashima Y, Kusakabe T, Kusakabe R, Terakita A, Shichida Y, Tsuda M. (2003). Origin of the vertebrate visual cycle: genes encoding retinal photoisomerase and two putative visual cycle proteins are expressed in whole brain of a primitive chordate. *J Comp Neurol*, 460: 180-90.

Nguyen V, Deschet K, Henrich T, Godet E, Joly JS, Wittbrodt J, Chourrout D, Bourrat F. (1999). Morphogenesis of the optic tectum in the medaka (*Oryzias latipes*): a morphological and molecular study, with special emphasis on cell proliferation. *J Comp Neurol*, 413: 385-404.

Nicol D and Meinertzhagen IA. (1988a). Development of the central nervous system of the larva of the ascidian, Ciona intestinalis. L.I. The early lineages of the neural plate. *Dev Biol*, 130: 721-36.

Nicol D and Meinertzhagen IA. (1988b). Development of the central nervous system of the larva of the ascidian, Ciona intestinalis. L.II. Neural plate morphogenesis and cell lineages during neurulation. *Dev Biol*, 130: 737-66.

Nicol D and Meinertzhagen IA. (1991). Cell counts and maps in the larval central nervous system of the ascidian Ciona intestinalis (L.). *J Comp Neurol*, 309: 415-29.

Nishida H. (1986). Cell division pattern during gastrulation of the ascidian, Halocynthia roretzi. *Dev Growth Differ*, 28: 191-201.

Nishida H. (1987). Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue restricted stage. *Dev Biol*, 121: 526-41

Nishida H. (1997). Cell fate specification by localized cytoplasmic determinants and cell interactions in ascidian embryos. *Int Rev Cytol*, 176: 245-306.

Nishida H and Satoh N. (1989). Determination and regulation in the pigment cell lineage of the ascidian embryo. *Dev. Biol*, 132: 355–367.

Ohtsuki H. (1991). Sensory organs in the cerebral vesicle of the ascidian larva, Aplidium sp.: an SEM study. *Zoolog Sci*, 8: 235-42.

Osborne PW, Benoit G, Laudet V, Schubert M, Ferrier DE. (2009). Differential regulation of ParaHox genes by retinoic acid in the invertebrate chordate amphioxus (*Branchiostoma floridae*). *Dev Biol*, 327: 252-62.

Parks CL, Banerjee S, Spector DJ. (1988). Organization of the transcriptional control region of the E1b gene of adenovirus type 5. *J Viro*, 62: 54-67.

Pei Z, Wang B, Chen G, Nagao M, Nakafuku M, Campbell K. (2011). Homeobox genes Gsx1 and Gsx2 differentially regulate telencephalic progenitor maturation. *Proc Natl Acad Sci USA*, 108: 1675-80.

Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, Benito-Gutiérrez EL, Dubchak I, Garcia-Fernàndez J, Gibson-Brown JJ, Grigoriev IV, Horton AC, de Jong PJ, Jurka J, Kapitonov VV, Kohara Y, Kuroki Y, Lindquist E, Lucas S, Osoegawa K, Pennacchio LA, Salamov AA, Satou Y, Sauka-Spengler T, Schmutz J, Shin-I T, Toyoda A, Bronner-Fraser M, Fujiyama A, Holland LZ, Holland PW, Satoh N, Rokhsar DS. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature*, 453: 1064-71.

Razy-Krajka F, Brown ER, Horie T, Callebert J, Sasakura Y, Joly JS, Kusakabe TG, Vernier P. (2012). Monoaminergic modulation of photoreception in ascidian: evidence for a proto-hypothalamo-retinal territory. *BMC Biol*, 10: 45.

Rhinn M, Picker A, Brand M. (2006). Global and local mechanisms of forebrain and midbrain patterning. *Curr Opin Neurobiol*, 16: 5-12.

Ristoratore F, Spagnuolo A, Aniello F, Branno M, Fabbrini F, Di Lauro R. (1999). Expression and functional analysis of Cititf1, an ascidian NK-2 class gene, suggest its role in endoderm development. *Development*, 126(22): 5149-59.

Sakurai D, Goda M, Kohmura Y, Horie T, Iwamoto H, Ohtsuki H, Tsuda M. (2004). The role of pigment cells in the brain of ascidian larva. *J Comp Neurol*, 475: 70-82.

Salvini-Plawen L and Mayr E. (1961). In: *Evolutionary biology*, vol. 10. (Hecht MK, Steere WC, and Wallace B, eds). New York: Plenum Press, 207-63.

Samadi L and Steiner G. (2010). Conservation of ParaHox gene's function in patterning of the digestive tract of the marine gastropod *Gibbula varia*. *BMC Dev Biol*, 10: 74.

Sasakura Y, Mita K, Ogura Y, Horie T. (2012). Ascidians as excellent chordate models for studying the development of the nervous system during embryogenesis and metamorphosis. *Develop Growth Differ*, 54: 420-37.

Sasakura Y, Oogai Y, Matsuoka T, Satoh N, Awazu S. (2007). Transposon mediated transgenesis in a marine invertebrate chordate: *Ciona intestinalis*. *Genome Biol*, 8, Suppl1: S3.

Sato S and Yamamoto H. (2001). Development of pigment cells in the brain of ascidian tadpole larvae: insights into the origins of vertebrate pigment cells. *Pigment Cell Res*, **14**: 428-36.

Sato S , Masuya H, Numakunai T, Satoh N, Ikeo K, Gojobori T, Tamura K, Ide H, Takeuchi T, Yamamoto H. (1997). Ascidian tyrosinase gene: its unique structure and expression in the developing brain. *Dev Dyn*, 208: 363-74.

Satoh N. (2001). Ascidian embryos as a model system to analyze expression and function of developmental genes. *Differentiation*, 68: 1-12.

Satoh N. (2003). The ascidian tadpole larva: comparative molecular development and genomics. *Nat Rev Genet*, 4: 285-95.

Satoh N, Satou Y, Davidson B, Levine M. (2003). *Ciona intestinalis*: an emerging model for whole-genome analyses. *Trends Genet*, 19: 376-81.

Satou Y, Imai KS, Levine M, Kohara Y, Rokhsar D, Satoh N. (2003). A genomewide survey of developmentally relevant genes in Ciona intestinalis. I. Genes for bHLH transcription factors. Dev Genes Evol, 231: 213-21.

Satou Y, Takatori N, Yamada L, Mochizuki Y, Hamaguchi M, Ishikawa H, Chiba S, Imai K, Kano S, Murakami SD, Nakayama A, Nishino A, Sasakura Y, Satoh G, Shimotori T, Shin-I T, Shoguchi E, Suzuki MM, Takada N, Utsumi N, Yoshida N, Saiga H, Kohara Y, Satoh N. (2001). Gene expression profiles in Ciona intestinalis tailbud embryos. *Development*, 128: 2893-904.

Schepers GE, Taesdale RD, Koopman, P. (2002). Twenty pairs of Sox: Extent, homology, and nomenclature of the mouse and human Sox transcription factor families. *Dev Cell*, 3: 167-70.

Schubert M, Escriva H, Xavier-Neto J, Laudet V. (2006). Amphioxus and tunicates as evolutionary model systems. *Trends Ecol Evol*, 21: 269-77.

Scott MP and Weiner AJ. (1984). Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila. *Proc Natl Acad Sci USA*, 81: 4115-9.

Seibert J and Urbach R. (2010). Role of en and novel interactions between msh, ind, and vnd in dorsoventral patterning of the Drosophila brain and ventral nerve cord. *Dev biol*, 46: 332-45.

Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539.

Singh G, Kaur S, Stock JL, Jenkins NA, Gilbert DJ, Copeland NG, Potter SS. (1991). Identification of 10 murine homeobox genes. *Proc Natl Acad Sci USA*, 88: 10706-10.

Skeath JB. (1998). The *Drosophila* EGF receptor controls the formation and specification of neuroblasts along the dorsal-ventral axis of the *Drosophila* embryo. *Development*, 125: 3301-12.

Smith ST and Jaynes JB. (1996). A conserved region of engrailed, shared among all en- gsc- Nk1-, Nk2 and msh-class homeoproteins, mediates active transcriptional repression in vivo. *Development*, 122: 3141–50.

Squarzoni P, Parveen F, Zanetti L. Ristoratore F and Spagnuolo A. (2011). FGF/MAPK/Ets signaling renders pigment cell precursors competent to respond to Wnt signal by directly controlling Ci-Tcf transcription. *Development*, 138: 1421-32.

Takahashi H, Mitani Y, Satoh G, Satoh N. (1999). Evolutionary alterations of the minimal promoter for notochord-specific Brachyury expression in ascidian embryos. *Development*, 126: 3725-34.

Takimoto N, Kusakabe T, Horie T, Miyamoto Y, Tsuda M. (2006). Origin of the vertebrate visual cycle: III. Distinct distribution of RPE65 and beta-carotene 15,15'-monooxygenase homologues in Ciona intestinalis. *Photochem Photobiol*, 82: 1468-74.

Taniguchi K and Nishida H. (2004). Tracing cell fate in brain formation during embryogenesis of the ascidian *Halocynthia roretzi*. *Dev. Growth Differ*, 46:163-80.

Toresson H and Campbell K. (2001). A role for Gsh1 in the developing striatum and olfactory bulb of Gsh2 mutant mice. *Development*, 128: 4769-80.

Toresson H, Potter SS, Campbell K. (2000). Genetic Control of dorsal-ventral identity in the telencephalon: opposing roles for Pax6 and Gsh2. *Development*, 127: 4361-71.

Treisman J, Gönczy P, Vashishtha M, Harris E, Desplan C. (1989). A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell*, 59: 553-62.

Treisman J, Harris E, Wilson D, Desplan C. (1992). The homeodomain: a new face for the helix-turn-helix? *Bioessays* 14: 145-50.

Tsuda M, Kawakami I, Shiraishi S. (2003a). Sensitization and habituation of the swimming behavior in ascidian larvae to light. *Zoolog Sci*, 20: 13-22.

Tsuda M, Sakurai D, Goda M. (2003b). Direct evidence for the role of pigment cells in the brain of ascidian larvae by laser ablation. *J Exp Biol*, 206: 1409-17.

Valerius MT, Li H, Stock JL, Weinstein M, Kaur S, Singh G, Potter SS. (1995). *Gsh-1*: a novel murine homeobox gene expressed in the central nervous system. *Dev Dyn*, 203: 337-51.

Vienne A and Pontarotti P. (2006). Metaphylogeny of 82 gene families sheds a new light on chordate evolution. *Int J Biol Sci*, 2: 32-7.

Von Ohlen T and Doe CQ. (2000). Convergence of Dorsal, Dpp, and Egfr signaling pathways subdivides the *Drosophila* neuroectoderm into three dorsalventral columns. *Dev Biol*, 224: 362-72.

Von Ohlen T and Moses C. (2009). Identification of Ind transcription activation and repression domains required for dorsoventral patterning of the CNS. *Mech Dev*, 126: 552-62.

Von Ohlen T, Moses C, Poulson W. (2009). Ind Represses *msh* Expression in the Intermediate Column of the *Drosophila* Neuroectoderm, Through Direct Interaction With Upstream Regulatory DNA. *Dev Dyn*, 238: 2735-44.

Von Ohlen T, Syu LJ, Mellerick DM. (2007a). Conserved properties of the *Drosophila* homeodomain protein, Ind. *Mech Dev*, 124: 925-34.

Von Ohlen TL, Harvey C, Panda M. (2007b). Identification of an upstream regulatory element reveals a novel requirement for Ind activity in maintaining *ind* expression. Mech Dev, 124: 230-6.

Wada H, Saiga H, Satoh N, Holland PW. (1998). Tripartite organization of the ancestral chordate brain and the antiquity of placodes: insights from ascidian *Pax-2/5/8*, *Hox* and *Otx* genes. *Development*, 125: 1113-22.

Wada S, Katsuyama Y, Yasugi S, Saiga H. (1995). Spatially and temporally regulated expression of the LIM class homeobox gene *Hrlim* suggests multiple distinct functions in development of the ascidian, Halocynthia roretzi. *Mech Dev*, 51: 115-26.

Wada S, Tokuoka M, Shoguchi E, Kobayashi K, Di Gregorio A, Spagnuolo A, Branno M., Kohara Y, Rokhsar D, Levine M, Saiga H, Satoh N, Satou Y. (2003). A genomewide survey of developmentally relevant genes in Ciona intestinalis: II. Genes for homeobox transcription factors. *Dev Genes Evol*, 213: 222– 334.

Wakimoto BT, Turner FR, Kaufman TC. (1984). Defects in embryogenesis in mutants associated with the antennapedia gene complex of Drosophila melanogaster. Dev Biol, 102: 147-72.

Walls GL. (1942). The Vertebrate Eye and its Adaptive Radiation. New York, Hafner.

Wegner M. (1999). From head to toes: The multiple facets of Sox proteins. *Nucleic Acids Res*, 27: 1409-20.

Wegner M. (2005). Secrets to a healthy Sox life: lessons for melanocytes. *Pigment Cell Res*, 18: 74-85.

Weiss JB, Von Ohlen T, Mellerick D, Dressler G, Doe CQ, Scott MP. (1998). Dorsoventral patterning in the *Drosophila* Central Nervous System: the intermediate neuroblasts defective Homeobox gene specifies intermediate column identity. Genes Dev, 12: 3591-602.

Wilson M and Koopman P. (2002). Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. *Curr Opin Gent Dev*, 12:441-6.

Winterbottom EF, Illes JC, Faas L, Isaacs HV. (2010). Conserved and novel roles for the Gsh2 transcription factor in primary neurogenesis. *Development* 137: 2623-31.

Winterbottom EF, Ramsbottom SA, Isaacs HV. (2011). Gsx transciption factor repress Iroquois gene expression. *Dev Dyn*, 240: 1422-9.

Wurst W and Bally-Cuif L. (2001). Neural plate patterning: upstream and downstream of the isthmic organizer. *Nat Rev Neurosci*, 2: 99-108.

Yamada L, Kobayashi K, Degnan B, Satoh N, Satou Y. (2003). A genomewide survey of developmentally relevant genes in *Ciona intestinalis*. IV. Genes for HMG transcriptional regulators, bZip and GATA/Gli/Zic/Snail. *Dev Genes Evol*, 213: 245-53.

Yoshida K and Saiga H. (2011). Repression of *Rx* gene on the left side of the sensory vesicle by Nodal signaling is crucial for right-sided formation of the ocellus photoreceptor in the development of *Ciona intestinalis*. *Dev biol*, 354: 144-50.

Yun K, Garel S, Fischman S, Rubenstein JL. (2003). Patterning of the lateral ganglionic eminence by the Gsh1 and Gsh2 homeobox genes regulates striatal and olfactory bulb histogenesis and the growth of axons through the basal ganglia. *J Comp Neurol*, 461: 151-65.

Yun K, Potter SS, Rubenstein JLR. (2001). Gsh2 and Pax6 play complementary roles in dorsoventral patterning of the mammalian telencephalon. *Development*, 128: 193-205.

Ringraziamenti...

La via che mi ha portato a poter finalmente scrivere queste righe conclusive non è stata né breve né semplice. In alcuni momenti mi è parsa perfino al di là delle mie possibilità, eppure eccomi qui, dall'altra parte di questa "avventura". È quindi questo il momento di ringraziare con tutto il cuore coloro che con la loro presenza, il loro affetto, la loro fiducia, mi hanno incoraggiata e sostenuta negli ultimi tre anni, e non solo.

Ringrazio Nietta, il mio tutor da sempre. Sono sbarcata quasi per caso nel suo laboratorio nei lontani tempi del tirocinio per la laurea triennale, e praticamente da allora muovo i miei passi nel mondo della ricerca sotto la sua guida costante. Se potrò continuare a lavorare nel complicato mondo della ricerca, quello che ho appreso alla Stazione Zoologica in questi anni sarà senza dubbio un significativo bagaglio da portare con me.

Desidero poi dedicare un enorme grazie a Margherita, Laura e Mena, insostituibili presenze a cui ricorrere non solo per questioni scientifiche, più simili a delle zie affettuose che non a dei "capi". Grazie per esserci state ogni volta che ho fatto capolino nel vostro studio!

Se l'atmosfera che ho respirato a lavoro in tutto questo tempo è stata serena e allegra, e se le pause caffè sono sempre state dei preziosi momenti di risate, confidenze e sfoghi, lo devo principalmente ad amiche e amici, di vecchia data o meno.

Voglio ringraziare in primo luogo Anna ed Elena per la loro amicizia incondizionata e per essere sempre lì per me, pronte ad aiutarmi, ascoltarmi e incoraggiarmi. Cara Anna, quello che ci lega non può essere brevemente spiegato con qualche parola... Ma mi sento libera di dirti solo: grazie Amica mia, perché so che leggendo queste parole capirai, come sempre, ciò che voglio dire. Grazie infinite anche a te, Elena, dolce/ruvida amica meravigliosa e...compagna di aperitivi! Sei una delle persone a cui tengo di più e una delle poche in cui ripongo la mia fiducia... Il rafforzamento della nostra amicizia è uno dei risvolti più positivi di questi anni all'Acquario.

Grazie anche a Mara e Leo, ottimi colleghi e amici, senza i quali i "tremolii" del mio sopracciglio o le battute sui livelli di pH nell'aria non sarebbero la stessa cosa.

Dedico un grosso grazie a Rosa, compagna di viaggio e punto di riferimento in questo complicato percorso, iniziato insieme un'eternità fa. Ci avresti creduto in quel lontano primo giorno di corsi che saremmo arrivate fin qui insieme?

Vorrei ringraziare Claudia, collega da tanto ma amica da poco. Galeotta fu Marsiglia e chi la scelse... Grazie per i confronti scientifici, per l'entusiasmo che trasmetti nel cercare di saperne sempre di più, di crescere e migliorarsi. Grazie anche per avermi dato la possibilità di conoscerti meglio. Al prossimo viaggio!

Vorrei poi ringraziare tutti i membri del gruppo di Biologia Cellulare e dello Sviluppo, e in particolare: Annamaria Locascio, Salvatore D'Aniello, Ina Arnone e Paolo Sordino per gli utili

suggerimenti e commenti al mio lavoro, fornitimi nel corso del tempo; Alessandro Amoroso per il supporto tecnico; Margherita Perillo, Rossella Annunziata e Carmen Andrikou per la disponibilità e la pazienza nel fornire reagenti e spiegazioni, e più in generale tutti gli altri "Arnonidi", colleghi sempre amichevoli e cordiali; Fabio, Luigi Caputi, Ashwani e Quirino, perché senza di loro il nostro gruppo non sarebbe più lo stesso; tutti gli studenti e i vari componenti del gruppo, passati ed attuali, che hanno contribuito a rendere ogni giorno di lavoro un po' meno faticoso e un po' più allegro.
Grazie a Giovanna Benvenuto per l'aiuto con il confocale e a Marco Borra per la pazienza e la professionalità con cui mi ha permesso di approcciarmi alle qPCR.

Ringrazio infine coloro che costituiscono il vero fulcro del mio mondo, gli incrollabili pilastri della mia vita, i sostegni sicuri di ogni mio passo. Mi riferisco naturalmente innanzitutto ai miei genitori e a mio fratello, che hanno sempre riposto in me fiducia e ammirazione, anche se ciascuno a suo modo.
E mi riferisco a Gianluca, l'Amore della mia vita. Ti ringrazio per tutto, per ogni singolo momento trascorso insieme, per avermi sempre sopportato e supportato (specialmente in questi tre anni di dottorato), per la pazienza e l'amore con cui sai starmi accanto come nessun altro potrebbe fare.
Ringrazio i miei suoceri, Alessandro, zio Carmine e zia Maria, che rappresentano per me una seconda vera famiglia, in grado di starmi vicino ed aiutarmi con dolcezza e affetto impagabili.

Ringrazio il Signore per il traguardo appena raggiunto, per avermi dato la forza di riuscirci e per le meravigliose persone di cui mi ha circondato per aiutarmi a raggiungerlo.